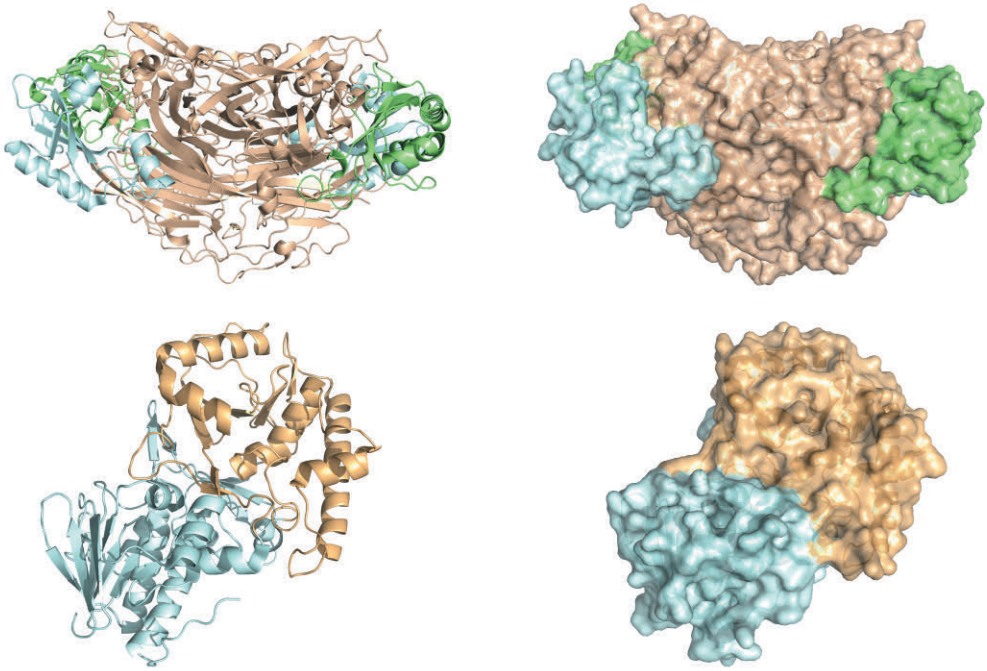
Åbo Akademi University

Leonor Lopes de Carvalho

Bioinformatics Characterization of Amine Oxidases



Leonor Lopes de Carvalho

was born on November 27, 1985 in Vila Franca de Xira, Portugal. In 2011 she obtained a Bachelor degree in Computer Science from the University of Gävle, Sweden. She graduated from Turku University in 2013 with a Master of Science in Bioinformatics. This PhD thesis project in Biochemistry has taken place during 2014-2018 under the supervision of Adjunct Professor Tiina Salminen at the Faculty of Science and Engineering.

Åbo Akademi University Press
Tavastgatan 13, FI-20500 Åbo, Finland
Tel. +358 (02) 215 3478
E-mail: forlaget@abo.fi

Sales and distribution:
Åbo Akademi University Library
Domkyrkogatan 2–4, FI-20500 Åbo, Finland
Tel. +358 (0)2 215 4190
E-mail: publikationer@abo.fi

Bioinformatics Characterization of Amine Oxidases

Leonor Lopes de Carvalho

Biochemistry
Faculty of Science and Engineering
Åbo Akademi University
Turku, Finland
2018

From the Faculty of Science and Engineering, Åbo Akademi University & National Doctoral Programme in Informational and Structural Biology

Supervised by:

Adjunct Professor Tiina A. Salminen
Faculty of Science and Engineering
Åbo Akademi University
Turku, Finland

Reviewed by:

Dr. Heidi Kidron
Faculty of Pharmacy
University of Helsinki
Helsinki, Finland

and

Docent Nina Hakulinen
Faculty of Science and Forestry
University of Easter Finland
Joensuu, Finland

Opponent:

Adjunct Professor Henri Xhaard
Faculty of Pharmacy
University of Helsinki
Helsinki, Finland

Cover, from left: Top: cartoon and surface representations of the crystal structure of hAOC3 (PDB ID 4BTX). Bottom: cartoon and surface representations of the homology model from SynPAO (publication IV). Image by Leonor Lopes de Carvalho.

ISBN 978-952-12-3762-1 (printed version)
ISBN 978-952-12-3763-8 (digital version)

Painosalama Oy - Turku, Finland 2018

Abstract

Amine oxidases catalyse the oxidative degradation of amines into the corresponding aldehyde, ammonia and hydrogen peroxide. They can be classified into different subfamilies based on the prosthetic group: (1) proteins with flavin adenine dinucleotide (FAD) make the FAD-dependent amine oxidase subfamily, which include monoamine oxidases (MAO) and polyamine oxidases (PAO) and (2) proteins with copper and quinocofactor make the other subfamily, which includes the copper-dependent amine oxidase (CAO) that have topaquinone (TPQ) as cofactor and the lysyl oxidases (LOX) that have the lysyl tyrosylquinone (LTQ) cofactor.

Most mammals have four CAO genes, each encoding a protein with a different function and substrate preference. The phylogenetic study with CAOs showed that these proteins can be classified into different subfamilies based on the conserved active site motif, X1-X2-Asn-Tyr-Asp. This classification is consistent with the different substrate preferences known for these proteins. Residue X2 can be used to distinguish AOC1 (X2=Tyr, diamine preference), AOC2 (X2=Gly, aromatic monoamine preference) and AOC3/4 (X2= Leu, aliphatic monoamine preference). The phylogenetic analysis led us to develop a novel classification system for the AOC3 and AOC4 proteins. These proteins can be classified based on the X1 residue from the motif, which is Leu in AOC3 and Met in AOC4.

When the N-glycosylations in hAOC1 were studied, it was found that the Asn110 site mainly contained oligomannosidic glycans. This site seems to be conserved in vertebrate CAOs and in hAOC1 it is located in a hydrophobic cleft on the surface. The attached glycans play a role in the correct folding, stabilization of the D2 domain, and protein secretion. Due to the high degree of conservation in the glycan composition, this N-glycosylation site likely plays a similar role in hAOC3 and porcine AOC1. The N-glycosylation sites at Asn168, Asn538 and Asn745 play a role in protein segregation and folding, and the Asn538 and Asn745 located near the inter-monomeric arms seem to be important for the dimer stability.

The interaction between hAOC3 and its leucocyte counter receptor, Siglec-9, is powerful tool for *in vivo* visualization of inflammation. In this thesis, the interactions between hAOC3 and

a Siglec-9 derived peptide were analysed taking into account for the first time the glycosylations in hAOC3. The peptide binds in the active site of hAOC3 competing for the same binding site as semicarbazide and imidazole. Moreover, the predicted binding mode of the peptide is in accordance with PET studies using rodent, rabbit and pig AOC3 proteins.

The polyamine oxidases (PAO) belong to the FAD-dependent amine oxidases. PAOs are involved in the catabolism of spermidine and their acetyl. The PAO from *Synechocystis sp.* PCC 6803 (SynPAO) was characterized in this thesis. SynPAO is involved in the polyamine back-conversion pathway, with spermidine being the preferred substrate. The structural analysis together with the phylogenetic tree revealed that Gln94, Tyr403 and Thr440 are predicted to be key residues in the active site to SynPAO.

Sammanfattning

Aminoxidaser katalyserar den oxidativa nedbrytningen av aminer till motsvarande aldehyd, ammoniak och väteperoxid. De kan klassificeras i olika underfamiljer baserat på den prostetiska gruppen: (1) proteiner med flavin-adenin-dinukleotid (FAD) utgör den FAD-beroende aminoxidas-underfamiljen med monoaminoxidaser (MAO) och polyaminoxidaser (PAO) och (2) proteiner med koppar och kino-kofaktor utgör den andra underfamiljen med kopparberoende aminoxidas (CAO) som har topakinon (TPQ) som kofaktor och lysyloxidaserna (LOX) som har lysyltyrosylkinon (LTQ).

De flesta däggdjur har fyra CAO-gener, som alla kodar för ett protein med specifik funktion och substratpreferens. Den fylogenetiska studien med CAO visade att dessa proteiner kan klassificeras i olika underfamiljer baserat på det konserverade aminosyramönstret X1-X2-Asn-Tyr-Asp i det aktiva stället. Denna klassificering överensstämmer med de olika substratpreferenser som är kända för dessa proteiner. Aminosyra X2 kan användas för att skilja AOC1 (X2 = Tyr, diaminpreferens), AOC2 (X2 =Gly, aromatisk monoaminpreferens) och AOC3/4 (X2 = Leu, alifatisk monoaminpreferens). På basen av den fylogenetiska analysen utvecklades ett nytt klassificeringssystem för AOC3- och AOC4-proteinerna. Dessa proteiner kan klassificeras baserat på X1-aminosyran från aminosyramönstret, som är Leu i AOC3 och Met i AOC4.

Studierna av N-glykosyleringarna i hAOC1 visade att Asn110-positionen huvudsakligen innehöll oligomannosidiska glykaner. Denna position verkar bevaras i vertebrat-CAO och i hAOC1 är den lokaliserad i en hydrofob klyfta på ytan. De bifogade glykanerna spelar en roll vid korrekt veckning, stabilisering av D2-domänen och proteinsekretion. På grund av den höga graden av konservation i glykankompositionen spelar denna N-glykosyleringposition sannolikt en liknande roll i hAOC3 och gris AOC1. N-glykosyleringspositionerna vid Asn168, Asn538 och Asn745 spelar en roll vid proteinsegregering och veckning, och Asn538 och Asn745, som är belägna nära de intermonomera armarna, verkar vara viktiga för dimerstabiliteten.

Samspelet mellan hAOC3 och dess leukocytbindningspartner Siglec-9 är ett kraftfullt verktyg för in vivo visualisering av inflammation. I denna avhandling analyserades

samverkningarna mellan hAOC3 och en peptid härledd från Siglec-9 och för första gången beaktades glykosyleringarna på hAOC3. Peptiden binder till det aktiva stället hos hAOC3 där den konkurrerar med semikarbazid och imidazol om samma bindningsställe. Det förutspådda bindningssättet för peptiden stämmer dessutom överens med PET-studier där AOC3 från gnagare, kanin och gris har använts.

Polyaminoxidaserna (PAO) hör till de FAD-beroende aminoxidaserna. PAO är involverade i katabolismen av spermidin och dess acetyl. PAO från Synechocystis sp. PCC 6803 (SynPAO) karakteriserades i denna avhandling. SynPAO är involverat i polyaminers exo-omvandling och föredrar spermidin som substrat. Den strukturella analysen tillsammans med det fylogenetiska trädet visade att Gln94, Tyr403 och Thr440 kan vara viktiga aminosyror i det aktiva stället hos SynPAO.

Table of contents

List of original publications

This thesis is based on original publications that are referred in the text by Roman numerals (I-IV). Publications are reprinted with permission from the publishers.

I. **Leonor Lopes de Carvalho**, Eva Bligt-Lindén, Ramaiah Arunchalam, Mark S. Johnson and Tiina A. Salminen. (2018) Evolution and functional classification of mammalian copper amine oxidases. Manuscript

II. Elisabeth Gludovcz, Daniel Maresch, **Leonor Lopes de Carvalho**, Verena Puxbaum, Laurenz J. Baier, Leander Sützl, Gabriela Guédez, Clemens Grünwald-Gruber, Barbara Ulm, Sophie Pils, Robin Ristl, Fredrich Altmann, Bernd Jilma, Tiina A. Salminen, Nicole Borth and Thomas Boehm. (2018) Oligomannosidic glycans at Asn110 are essential for secretion of human diamine oxidase. Journal Biological Chemistry 293(3): 1070-1087.

III. **Leonor Lopes de Carvalho**, Heli Elovaara, Jerôme de Ruyck, Gerard Vergoten, Sirpa Jalkanen, Gabriela Guédez and Tiina A. Salminen. (2018) Mapping the interaction site and effect of the Siglec-9 inflammation biomarker on human primary amine oxidase. Scientific Reports 8(1): 2086.

IV. Khanittha Samasil, **Leonor Lopes de Carvalho**, Pirkko Mäenpää, Tiina A. Salminen, Aran Incharoensakdi. (2017) Biochemical characterization and homology modeling of polyamine oxidase from cyanobacterium Synechocystis sp. PCC 6803. Plant Physiology and Biochemistry 119: 159-169.

Contributions of the author

The author and supervisor designed of the work and wrote the manuscripts together. All co-authors contributed to the final manuscripts.

I. The author did the sequence alignments, phylogenetic study, modelling and data analysis and mainly wrote the manuscript.

II. The model of porcine AOC1 was made by the author, the structural comparison of the N-glycosylations sites in human AOC1 and AOC3 crystal structures, and the model of porcine AOC1 (pAOC1) were also performed by the author. The author wrote the sections regarding her work.

III. The author analysed, evaluated and interpreted the computational work and wrote the corresponding parts.

IV. The author performed the phylogenetic study and the molecular docking. The author also helped in analysing the results and wrote the sections regarding her work.

Acknowledgments

This thesis work was carried out in the Structural Bioinformatics Laboratory (SBL) in the Faculty of Science and Engineering, at Åbo Akademi University in Turku during 2014-2018. During the years many people took part of my thesis journey and I would like to thank everyone that made this thesis possible.

First I would like to thank *Adjunct Professor Tiina A. Salminen* and *Professor Mark Johnson*, for the opportunity to do my thesis at SBL and also for the excellent computing facilities provided. I would like especially to thank my supervisor *Adjunct Professor Tiina A. Salminen* for giving me the opportunity to do this thesis in her research group. Thank you for sharing your knowledge and for your guidance. I would like to thank *Dr. Heidi Kidron* and *Docent Nina Hakulinen* for taking time to read my thesis and for the valuable comments that were used to improve my thesis. I would like to thank all the co-authors and the collaborators in the research groups of *Academy Professor Sirpa Jalkanen*, *Dr. Thomas Boehm*, *Dr. Jérôme de Ruyck* and *Professor Gerard Vergoten*.

I would also like to thank the members of my thesis committee, *Professor (p.t.) Outi Salo-Ahen* and *Docent Lari Lehtiö* for all the discussions and feedback on my thesis work. I would like to acknowledge my graduate school "National Doctoral Programme in Informational and Structural Biology (ISB)", its director *Professor Mark Johnson* and coordinator *Fredrik Karlsson*. I am glad I was part of this outstanding graduate school, with all its events and scientific meetings.

I would also like to thank all the past and present members of SBL for creating a good work atmosphere in the office, for all the company in the lunch and coffee breaks and for all the interesting scientific and non-scientific discussions. I would like to thank specially *Fredrik Karlsson* for all the help with practical matters and *Jukka Lehtonen* for saving me when I had computer problems. I would like to thank *Käthe Dahlström* for all the fun times we had while traveling, for the shared wine, macaroons, and "bajete", and for all the support during my times at SBL. I want to thank *Mia Åstrand* for being such a great colleague, for all the very interesting conversations we had in our lunch and coffee breaks, and for reading my mind when its time to go for coffee. I want to thank

Gabriela Guédez for the nice collaboration and co-authorship. A very special thanks also goes to *Polytimi Dimitirou*, who has been a colleague and friend since my master degree, I hope we continue to share the surrealism in our lives for the years to come. I would also like to thank *Tomi Airenne, Mahlet Tamirat, Vipin Ranga, Nitin Agrawal* and *Ida Alanko* for creating a nice work environment and for the nice conversations we had during the years.

I would also like to thank my friends outside of academia that reminded me that life is more than looking at protein structures. *Inês*, thank you for being such a good friend and the best roommate someone can ask. *Lena, Tiago* and *Tuire* (and *Afonso)* thank you for bringing a piece of Portugal to Finland. I want to thank my friends *Sara* and *Sandra*; although you are very far away it still feels that you are near by, thank you for all the support.

I want to thank my family for the encouragement and help they gave me through these years. A particular thank you goes to my grandparents *José* and *Maria* and to my mother *Rita* that always encouraged me to study and never give up. I want also to acknowledge *Juri* and *Viktoria*, thank you for the support and for making me feel at home. I saved the best for last, *Mikko* thank you for being my rock during the hard times of my doctoral studies, thank you for all the shared the good times, for encouraging me to become a better version of myself and for showing me the real meaning of love and companionship. This thesis wouldn't be possible without you.

Thank you also to the several funding institutions that helped me thorough my doctoral studies. I acknowledge ISB, Turku Center for System Biology, Medicinska Understödsföreningen Liv och Halsa r.f., Tor, Joe och Pentti Borgs Minnesfond, Finnish Cultural Foundation, and Åbo Akademi for their financial support.

Leonor Lopes de Carvalho

Turku, October 2018

Abbreviations

2HP	2-hydrazinopyridine
3D	Three-dimensional
AO	Amine oxidase
AOC1	Copper amine oxidase 1; alternative names: diamine oxidase (DAO), histaminase
AOC2	Copper amine oxidase 2; alternative names: retina-specific amine oxidase (RAO)
AOC3	Copper amine oxidase 3; alternative names: vascular adhesion protein-1 (VAP-1)
AOC4	Copper amine oxidase 4; alternative names: soluble amine oxidase (SAO)
AGAO	*Arthrobacter globiformis* amine oxidase
AtPAO	*Arabidopsis thaliana* polyamine oxidase
BC	Back conversion
BLAST	Basic Local Alignment Search Tool
CAO	Copper amine oxidase
CNX	Calnexin
CRT	Calreticulin
ECAO	*Escherichia coli* amine oxidase
ER	Endoplasmic reticulum
ERAD	Endoplasmic reticulum-associated protein degradation
FAD	Flavin adenine dinucleotide
Glc	D-Glucose
GlCNAc	N-acetylglucosamine
GnT1	N-acetyl-glucosaminyltransferase 1
HPAO	*Hansenula polymorpha* amine oxidase
HMM	Hidden Markov model
HvPAO	*Hordeum vulgare* polyamine oxidase
JTT	Jones Taylor Thornton
LG	Le Gascuel
LOX	Lysyl oxidase
Man	D-Mannose
MAO	Monoamine oxidase
ML	Maximum likelihood
Nac	D-glucosamine
NCBI	National Center for Biotechnology Information
NJ	Neighbour joining
NMR	Nuclear Magnetic Resonance
OsPAO	*Oryza sativa* polyamine oxidase
PAO	Polyamine oxidase
PCA	Principal component analysis
PDB	Protein Data Bank
PET	Positron emission tomography
PPLO	*Pichia pastoris* amine oxidase
PSAO	*Pisum sativum* amine oxidase
Siglec	Sialic acid binding Immunoglobulin-like lectins
SMO	Spermine oxidase
SynPAO	*Synechocystis sp.* PCC 6803 polyamine oxidase
TC	Terminal catabolism
TPQ	2,4,5-trihydroxyphenylalanine quinone, topaquinone
ZmPAO	*Zea mays* polyamine oxidase

1. Introduction

Proteins are the engines of life. They are macromolecules that play a variety of roles in supporting all living organisms. There are twenty different amino acids that build the proteins and the genetic information stored in the DNA determines how the amino acids are assembled. The polypeptide chain of amino acids carries information that specifies how proteins fold into their three-dimensional (3D) structure. For their function, proteins may also require cofactors that can be e.g. an inorganic group or an amino acid that has undergone a posttranslational modification. The 3D structure of a protein determines its function in which its ability to interact with other molecules or with other proteins has an important role. For example, proteins can function as hormones that transmit signals to coordinate biological processes (e.g. growth hormone) (Greenwood and Landon, 1966), they can help in the immune system by protecting our body against virus and bacteria (e.g. immunoglobulin G) (Mallery et al., 2010), they can have a structural role (e.g. actin) (Doherty and McMahon, 2008) and they also can work as enzymes (e.g. amine oxidases) (Mondovì and Agrò, 1982). In general, enzymes are biocatalysts, which catalyse a specific reaction, and they can be classified as transferases, hydrolases, lyases, isomerases, ligases and oxidoreductases according to the catalysed reaction. For example, amine oxidases, which belong to oxidoreductases catalyse a variety of biological amines to their corresponding aldehyde and each amine oxidase subfamily has a preference for a certain amine.

In the 1960s decade computers emerged to be an important tool in biological research (Hagen, 2000). The increasing number of available amino and nucleic acid sequences and the idea that macromolecules carry information together with high-speed computers united biology and information technology (Hagen, 2000; Luscombe et al., 2001), to create bioinformatics. Bioinformatics aims to organize data allowing researchers to easily access data, to develop tools and resources for data analysis and to use those tools to analyse the data in a biologically meaningful way (Luscombe et

al., 2001). Based on the source of information, bioinformatics can be divided into two classes: sequence based and structure based (Choong et al., 2013). Data analysis of genomes, sequence alignments, phylogenetic analysis and identification of motifs are examples of sequence based bioinformatics (Choong et al., 2013; Luscombe et al., 2001). Structure based bioinformatics is related to the analysis of 3D structure of proteins, structure based alignments, molecular simulations (docking predictions and molecular dynamics) and intermolecular interactions (Choong et al., 2013; Luscombe et al., 2001). The two classes of bioinformatics are closely linked and commonly used in an integrative manner.

Combining the 3D structure of a protein (X-ray crystallography, NMR and homology modelling) with data about the protein's function (experimental data and different databases), occurrence in different genomes (homologous sequences, sequence alignments and phylogenetic analyses) and the interactions with other molecules (protein-protein interactions, docking simulations and molecular dynamics) integrates the information and puts in it a meaningful biological context (Gerstein, 2000). The work described throughout this thesis explores an integrative data analysis approach. Structural and sequence data, functional information, genome occurrence and molecular interactions were used together to characterize mammalian copper amine oxidases and one cyanobacteria polyamine oxidase.

2. Review of the literature

2.1. Amine oxidases

Amine oxidases (AO) are a family of proteins found in all living organisms, they are diverse in terms of structure, catalytic and substrate oxidation mechanisms (Toninello et al., 2006). AO catalyse the oxidative degradation of amines into the corresponding aldehyde, ammonia and hydrogen peroxide (Figure 1). These proteins can be classified into two distinct classes based on their prosthetic group: (1) the flavin adenine dinucleotide (FAD)-dependent AOs and (2) the copper-dependent AOs.

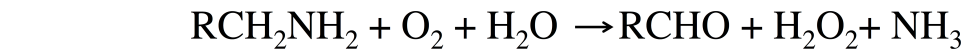


Figure 1. General reaction catalysed by amine oxidases.

The FAD-dependent AOs include two mitochondrial monoamine oxidases (MAO A and MAO B; EC 1.4.3.4) and the cytosolic polyamine oxidase (PAO; EC 1.5.3.17). MAOs metabolize neurotransmitters and MAO inhibitors are used to treat depression, anxiety disorders, Parkinson and Alzheimer´s disease (Yamada and Yasuhara, 2004). PAOs are involved in the catabolism of spermine, spermidine and their acetyl derivatives at their secondary amino group (Cona et al., 2006).

The copper-dependent AOs can be further classified based on their quinone cofactor. In the copper amine oxidases (CAOs, EC 1.4.3.21 and EC 1.4.2.22) the cofactor is 2,4,5-trihydroxyphenylalaninequinone (TPQ), whereas in the lysyl oxidases (LOXs, EC 1.4.3.13) the cofactor is lysine tyrosylquinone (LTQ) (Klinman, 1996). The LOX enzymes convert lysine and hydrxyolysine residues in collagens and elastin into highly reactive aldehydes, having a role in the remodelling of the extracellular matrix (Grau-Bové et al., 2015). Humans have five members of the LOX family, which are divided into two subgroups based on the N-terminal domain: the LOX and LOX1 form one subgroup and LOX2, LOX3 and LOX4 the other subgroup (Finney et al., 2014). CAOs are divided into several subclasses with different tissue and substrate

preference and they oxidize the primary amino group of the amine substrates (Cona et al., 2006; Finney et al., 2014). This thesis focuses on two AO families, CAOs and PAOs, which are described in detail in the following sections.

2.2. Polyamine oxidases

2.2.1. Function and catalytic reaction

Polyamines are positively charged molecules present in all prokaryotic and eukaryotic cells. The most common polyamines are putrescine, spermidine and spermine. They are involved in cell growth, development and adaptation against environmental stress (Bouchereau et al., 1999; Cona et al., 2006; Tavladoraki et al., 2016). One of the potential uses of PAOs is as biosensors, to detect biogenetic amines for food analysis and human diagnostics (Boffi et al., 2015). PAOs are FAD-dependent AO that catalyse the oxidation of polyamines via an oxidative cleavage of the α-CH bond of the substrate to form an imine product with associated reduction of the non-covalently bound FAD cofactor (Figure 2A). The imine product is non-enzymatically hydrolysed to the corresponding aldehyde and ammonia or amine (for secondary or tertiary amine substrates), the reduced FAD cofactor reacts with oxygen forming hydrogen peroxide and re-oxidizing the cofactor, completing the catalytic cycle (Binda et al., 2001) (Figure 2A). The differences in the reaction products reflect the differences in the position and orientation of the substrate inside the catalytic site.

In plants, two types of PAOs have been identified: (1) proteins that oxidase the carbon on the *endo*-site of the N^5-nitrogen (terminal catabolic pathway) and (2) proteins that oxidize the carbon on the *exo*-site of the N^5-nitrogen (back-conversion pathway) (Figure 2B). PAO from *Zea mays* (maize; ZmPAO), an *Oryza sativa* PAO (rice; OsPAO7) and two *Hordeum vulgare* PAOs (barley; HvPAO1 and HvPAO2) are involved in the terminal catabolic pathway (Cervelli et al., 2006; Liu et al., 2014a; Tavladoraki et al., 1998). They oxidize spermidine and spermine to produce 4-aminobutanal and N-(3-aminopropyl)-4-aminobutanal, respectively (Figure 2B). *Arabidopsis thaliana* PAOs (AtPAO1-5), several *O. sativa* PAOs (OsPAO1,

OsAO4-5), animal and yeast PAOs are involved in the back-conversion pathway (Ahou et al., 2014; Cervelli et al., 2015; Cona et al., 2006; Fincato et al., 2011; Liu et al., 2014a; Ono et al., 2012). These proteins oxidize spermine to spermidine and/or spermidine to putrescine (Figure 2B).

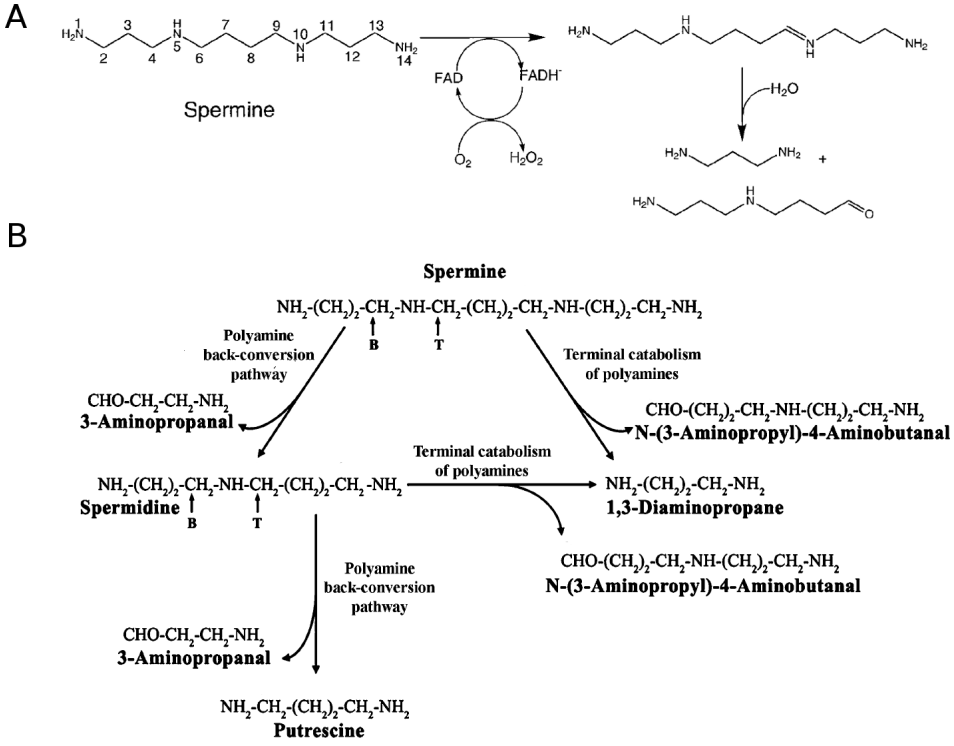


Figure 2. A: ZmPAO reaction scheme for spermine oxidation (adapted from Binda et al., 2001). B: Schematic representation of spermine and spermidine oxidation by PAOs. Arrowheads B or T indicate the oxidised carbon atoms in the back conversion or terminal conversion, respectively (adapted from Fincato et al., 2011).

2.2.2. The PAO fold

The 3D structure of the ZmPAO is solved both in native state (Binda et al., 1999) and in complex with spermidine and spermine (Fiorillo et al., 2011). The structure of ZmPAO is composed of two domains, a FAD-binding domain and a substrate-binding domain (Figure 3A). The interface between the domains forms a 30Å long U-shaped

tunnel that defines the active centre (Binda et al., 1999) (Figure 3B). The structure of yeast PAO (Fms1) is also known in the native form and in complex with spermine (Huang et al., 2005).

Despite having only 20% sequence identity, Fms1 and ZmPAO have a similar fold. Although both proteins have several common features, their structures revealed important differences (Fiorillo et al., 2011) e.g. the isoaloxazine ring of the cofactor is planar in Fms1 but bent in ZmPAO and some of the key residues in the active site are not totally conserved. As a result of these differences, Fms1 has a more hydrophobic substrate-binding site compared to that of ZmPAO. The binding of spermine is also different in both proteins. In Fms1 spermine binds more shallowly in the tunnel, with the C^6 atom too far from the N^5 of FAD to allow the production of 3-(aminopropyl)4-aminobutyraldehyde and 1,3-diaminopropane, but the C^4 is close enough to N^5 of FAD to produce spermidine and 3-aminopropanal. In ZmPAO, spermine binds deep inside of the tunnel and the C^6 atom is positioned closer to FAD whereas the C^4 is further away. These differences are responsible for the differences between the *endo*- and *exo*-mode of oxidation.

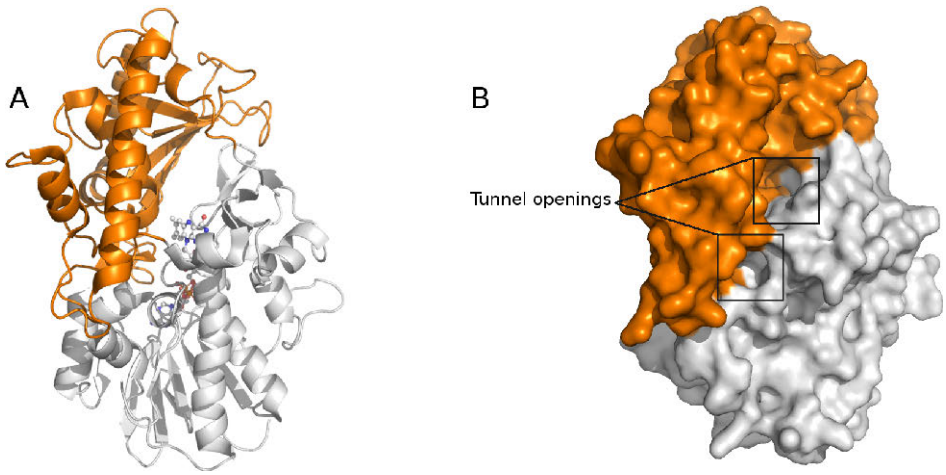


Figure 3. A: The PAO fold, orange ribbons represent the catalytic domain and white ribbons the FAD binding domain. FAD cofactor is shown as white ball and stick representation. B: Surface view showing the entrances of the U-shaped active site tunnel highlighted with boxes with the PAO fold in orange and §the catalytic domain in white colour.

2.3. Copper amine oxidases

Humans have three CAOs: (1) human diamine oxidase (AOC1), (2) retina-specific CAO (AOC2) (3) human vascular adhesion protein-1 (AOC3). Some mammals also contain an additional soluble CAO (AOC4), however, humans have a non-functional AOC4 protein due to a stop codon (Schwelberger, 2010). Rodents are also missing the AOC4 protein, since they only have small fragments of the *AOC4* gene (Schwelberger, 2007). Moreover, rat is missing the *AOC2* gene (Zhang et al., 2003). In general, organisms with high levels of plasma CAO have the *AOC4* gene (Schwelberger, 2010). Despite having highly similar structures (reviewed in 2.3.5) all mammalian CAOs have different function and different substrate preferences (Bonaiuto et al., 2010; Finney et al., 2014).

2.3.1. AOC1

Human AOC1 is a soluble protein involved in the degradation of exogenous histamine mainly expressed in kidney, placenta, intestine and seminal vesicles (Elmore et al., 2002). It has been suggested to have functions related with cell proliferation, inflammation, allergic response and ischemia (McGrath et al., 2009). Decreased levels of hAOC1 have been linked with histamine intolerance (Maintz and Novak, 2007). There is also evidence that some pharmaceuticals might interfere with normal AOC1 function and cause harmful side effects (Maintz and Novak, 2007; McGrath et al., 2009; Sattler et al., 1985). The plasma concentrations of AOC1 increase several hundred-fold during pregnancy and low AOC1 concentrations increase the risk of premature pregnancy termination by 17-fold (Maintz et al., 2008). AOC1 has preference for diamines such as histamine and its methylated form (1-methylhistamine) (Elmore et al., 2002).

2.3.2. AOC2

Imamura and co-workers cloned AOC2 from human retina in 1997 (Imamura et al., 1997). To date a lot remains to be discovered about this protein, which is the least studied human CAO. AOC2 mRNA has been found in different tissues (lung, brain, kidney, cartilage, tonsil and heart) but enzyme activity was only found in retina

(Kaitaniemi et al., 2009). AOC2 preferably oxidises aromatic monoamines *in vitro* and it has shown activity towards 2-phenylethyamine, tryptamine, p-tyramine (Kaitaniemi et al., 2009).

2.3.3. AOC3

The AOC3 protein is primarily expressed under inflammatory conditions on the endothelial cell surface but it is also found in muscle cells and adipocytes (Jaakkola et al., 1999; Salmi et al., 1993; Zorzano et al., 2003). It is a dual function protein, working as an amine oxidase and as an adhesion molecule involved in leucocyte trafficking (Salmi and Jalkanen, 2001). Both functions are linked (Noonan et al., 2013), as the amine oxidase reaction products change the expression of some endothelial selectins involved in the leucocyte extravasation cascade (Jalkanen et al., 2007). In adipocytes, the suggested roles for AOC3 are related to glucose metabolism (Boomsma et al., 2005) and defence against subcutaneous bacterial infiltration (Shen et al., 2012). AOC3 activity in adipocytes has been shown to be down regulated by hypoxia (Repessé et al., 2015). AOC3 is a membrane bound protein, but it can be proteolytically cleaved to a soluble form in the blood (sAOC3, serum AOC3) (Abella et al., 2004; Kurkijärvi et al., 1998; Stolen et al., 2004). The substrates of AOC3 include the physiological amines methylamine and aminoacetone and benzylamine functions as an *in vitro* substrate (Shen et al., 2012). Recently, the leukocyte surface proteins Siglec (sialic acid-binding Immunoglobulin (Ig)-like lectin) -9 and -10 were identified as a counter-receptors for human AOC3 (hAOC3) (Aalto et al., 2011; Kivi et al., 2009). Moreover, the interaction mechanism between hAOC3 and Siglec-9 has been characterized (Aalto et al., 2011; Elovaara et al., 2016).

2.3.4. AOC4

AOC4 is a soluble plasma protein expressed in the liver and secreted into the bloodstream, making it the major serum amine oxidase in some mammalian species (Schwelberger, 2006). Bovine AOC4 (bAOC4) is the most studied AOC4 protein. The physiological polyamines spermidine and spermine, long hydrophobic amines and also benzylamine with hydrophobic substituents are the

preferred substrates for AOC4 (Bonaiuto et al., 2010; Di Paolo et al., 2003).

2.4. TPQ and the reaction mechanisms of CAOs

CAOs catalyse two different reactions: (1) an autocatalytic reaction that gives origin to TPQ and requires molecular oxygen and copper to occur (Figure 4A) and (2) the catalytic reaction converting primary amines to the corresponding aldehydes (Figure 4B). The catalytic reaction is divided into the reductive half-reaction and the oxidative half-reactions. TPQ, which was first identified by Janes et al. in 1990 (Janes et al., 1990) derives from a post-translationally modified tyrosine (Mure, 2004). The modified Tyr is part of a specific motif, Thr-X1-X2-Asn-Tyr-Asp (Parsons et al., 1995), hereafter referred as the active site motif. The relevancy of the conserved residues in the motif has been assessed by mutational studies (Choi et al., 1996). TPQ is buried inside the active site near the copper-binding site. It is quite mobile as it can flip around the $\chi 2$ and $\chi 1$ angles interchanging between the active (off-copper) conformation, where the O^5 atom points towards a conserved active site base (Asp), and the inactive (on-copper) conformation, where the O^5 is in direct contact with copper (Dawkes, 2001) (Figure 4C).

The two half-reactions of the CAO catalytic reaction follow a classical ping-pong mechanism (Figure 4B). In the reductive half-reaction, the amine substrate reacts with O^5 atom of the oxidized TPQ forming the substrate Schiff base. The catalytic base abstracts the Cα-proton from the substrate forming the product Schiff base (Mure et al., 2002). The product Schiff base is hydrolysed releasing aldehyde and leaving the enzyme reduced (Klema and Wilmot, 2012). In the oxidative half-reaction, O_2 oxidizes the reduced enzyme, releasing ammonia and hydrogen peroxide (Mure et al., 2002).

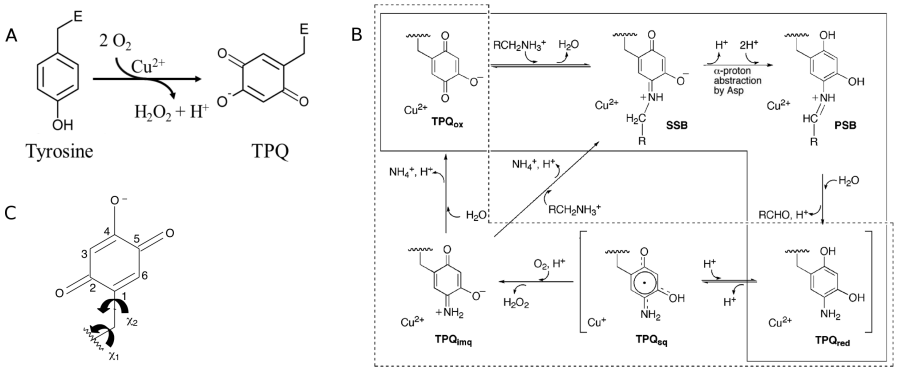


Figure 4. A: TPQ Biogenesis (Klema and Wilmot, 2012). B: Scheme showing the reaction mechanism of CAOs. The oxidative half-reaction is shown in dotted box and the reductive in solid box. TPQox, oxidized form of TPQ; SSB, substrate Schiff base intermediate; PSB, product Schiff base intermediate; TPQred, reduced form of TPQ; TPQsq, semiquinone form of TPQ; TPQimq, iminoquinone form of TPQ (adapted from (Finney et al., 2014; Mure, 2004)). C: Chemical structure of TPQ showing the torsion angles $\chi 1$ and $\chi 2$ (Dawkes, 2001).

2.5. The CAO fold

To date the structures of CAOs from different sources have been solved, which include the mammalian sources human and bovine (Table 1) and, despite the overall low sequence identity, they all adopt the same archetypal fold. The CAO fold is a heart shaped homodimer with each monomer having a ~70-80 kDa mass. CAOs from eukaryotic sources have additional mass due to glycosylation (Klema and Wilmot, 2012). Each monomer is composed of the D2-D4 domains (Figure 5A). The D2 and D3 domains have a similar α/β fold. Based on their high sequence identity in *Escherichia coli* (ECAO), they have presumably arisen via a gene duplication event (Parsons et al., 1995). The D4 domain, where the active site is located, is the largest and the most conserved domain (Figure 5A). The architecture of D4 is somewhat complex as there are two/three β-hairpins that protrude from one monomer to the other monomer. The active site is formed partially by residues belonging to the other monomer (Figure 5B), which has led to the speculation of a possible cooperation between the two monomers (Frébort et al., 1996). Recently, it was however found that there is no communication between the active sites in *Arthrobacter globiforis* CAO (AGAO)(Gaule et al., 2015). Moreover, the hairpin arms have been

thought to be important for the maintenance of the dimer but also to have a role in regulating substrate specificity (Klema and Wilmot, 2012; Wilce et al., 1997). The active site is buried in the end of a channel and it contains TPQ, a copper ion coordinated by three histidines and a conserved catalytic base (Asp) (Figure 5C). Depending on the source, the active site channel has different shape and properties. For example, in rodent and primate AOC3 the channel has different residues and shape (Bligt-Lindén et al., 2012). Moreover, due to these residue differences, some inhibitors might not bind equally well to rodent enzymes, which is problematic for the preclinical testing of these inhibitors (Bligt-Lindén et al., 2013).

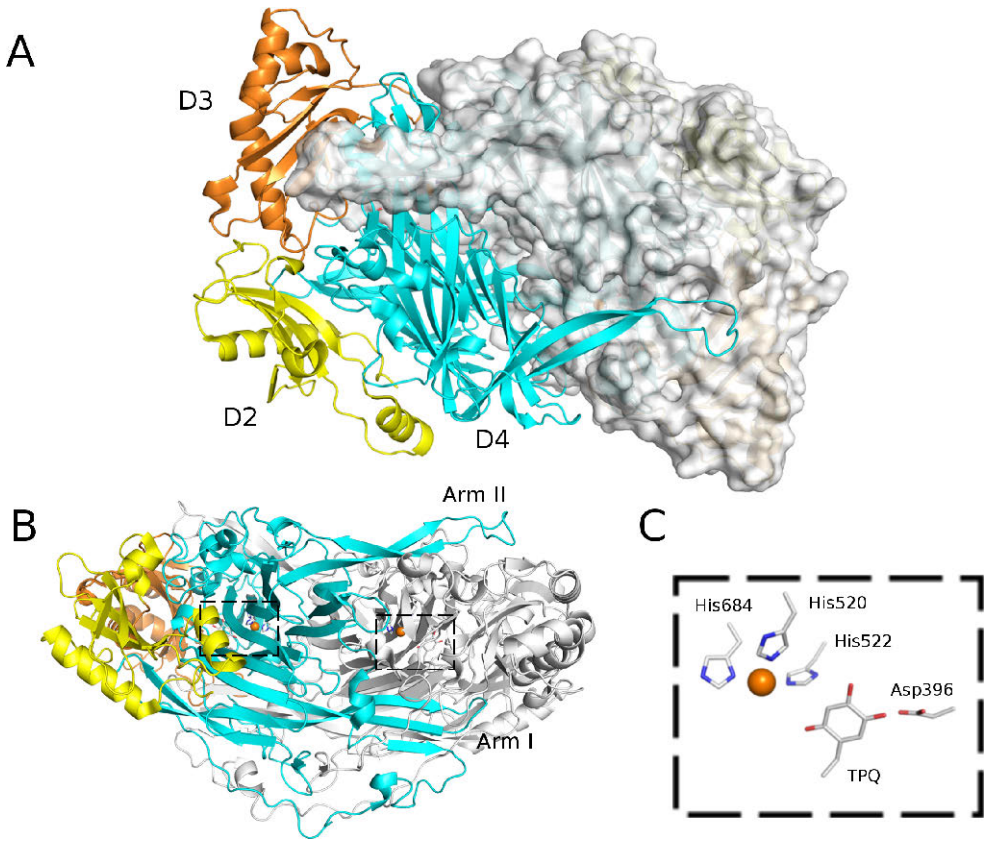


Figure 5. A: Top view of the CAO fold. The domains are color-coded in the monomer shown as cartoon where the other monomer of the dimer is shown as

grey surface. B: Side view of the CAO fold. Arm I from one monomer protruding near the active site of the other monomer. The catalytic centres are indicated with dashed boxes. C: Conserved catalytic centre in all CAOs with hAOC3 residue numbering. The images were prepared using the crystal structure of hAOC3 PDB ID 4BTY (Bligt-Lindén et al., 2013).

Table 1. Summary of the available crystal structures of mammalian CAOs in the Protein Data Bank (PDB).

Protein	Organism	Ligand	TPQ conformation	PDB ID	Reference
AOC1					
	Human		on-copper	3HI7	(McGrath et al., 2009)
	Human	Berenil	on-copper	3HIG	(McGrath et al., 2009)
	Human	Pentamidine	on-copper	3HII	(McGrath et al., 2009)
	Human		on-copper	3K5T	(McGrath et al., 2010)
	Human	Aminoguanidine	off-copper	3MPH	(McGrath et al., 2010)
AOC3					
	Human		on-copper	1PU4	(Airenne et al., 2005)
	Human		on-copper	1US1	(Airenne et al., 2005)
	Human		off-copper	2C10	(Jakobsson et al., 2005)
	Human	2-hydrazinopyridine	off-copper	2C11	(Jakobsson et al., 2005)
	Human		off-copper	3ALA	(Ernberg et al., 2010)
	Human	Imidazole	on-copper	2Y73	(Elovaara et al., 2011a)
	Human	Imidazole	off-copper	2Y74	(Elovaara et al., 2011a)
	Human	Pyridazinone	off-copper	4BTX	(Bligt-Lindén et al., 2013)
	Human	Pyridazinone	off-copper	4BTY	(Bligt-Lindén et al., 2013)
AOC4					
	Bovine		off-copper	1TU5	(Lunelli et al., 2005)
	Bovine	Clonidine	off-copper	2PNC	(Holt et al., 2008)

2.6. N-glycosylations in hAOC1 and hAOC3

The addition of glycans on a protein can dramatically alter the structure of the protein and consequently its function (Imperiali and O'Connor, 1999). For example, the carbohydrates covalently linked

to Asn residues, the N-glycans, play a key role in many biological processes (Eklund and Freeze, 2005). The N-glycans can have an affect on secretion, stability, expression, adhesion, signalling and folding of proteins (Bieberich, 2014; Tannous et al., 2015). The attachment of N-glycans to Asn residues may occur at the Asn-X-Ser/Thr consensus sequence (X can be any residue except Pro) on the protein (Tannous et al., 2015) (Figure 6).

In the ER, the Asn is modified with the attachment of pre-formed oligosaccharides made of three glucoses, nine mannoses and two N-acetylglucosamines. The first terminal glucose is trimmed by alpha-glucosidase I and the second glucose by alpha-glucosidase II (Figure 6). The two lectin chaperones, calnexin (CNX) and calreticulin (CRT), facilitate the folding. The immature proteins are re-glycosylated by glycosyltransferance (UGT1), which serves as a folding sensor (Tannous et al., 2015). The immature proteins re-enter the CNX/CRT cycle. The misfolded proteins or proteins recognized as aberrant products undergo the ER-associated degradation (ERAD) but the correctly folded proteins are exported from the ER to the Golgi, where the glycans are further processed (Bieberich, 2014).

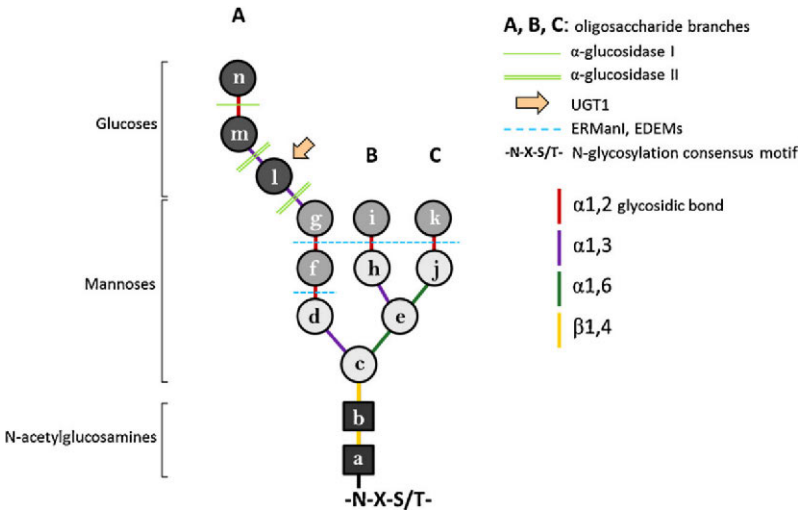


Figure 6. The N-glycan attached to the Asn-X-Ser/Thr motif. The pre-formed oligosaccharide is made of three glucoses (dark gray circles), nine mannoses (light gray circles) and two N-acetylglucosamines (black squares). The glycan processing

enzymes and their actions points are listed in the top right corner of the figure (Adapted from Tannous et al. (2015)).

The human AOC1 and AOC3 are both N-glycosylated and the crystal structures show that the N-glycans are distributed on the surface of these proteins (Figure 7). hAOC1 has four N-glycosylation sites, Asn110, Asn168, Asn538 and Asn745 (McGrath et al., 2009). Whereas hAOC3 has six N-glycosylated sites: Asn137, Asn232, Asn294, Asn592, Asn618 and Asn666 (Airenne et al., 2005; Maula et al., 2005). The composition of the glycan structures of hAOC1 has been recently resolved (Gludovacz et al., 2016). hAOC1^{Asn110} is occupied mainly by oligomannosidic glycans, hAOC1^{Asn168} by bi- or tetra-antennary complex glycans, hAOC1^{Asn538} and hAOC1^{Asn745} by bi-antennary complex type glycans (Gludovacz et al., 2016). Regarding hAOC3, only the composition of the glycan attached to Asn137 is known (Gludovacz et al., 2018). The mutation of the N-glycosylation sites in hAOC1 had an effect on protein secretion in particular hAOC1^{Asn110}, which is occupied by oligomannosidic glycans (Gludovacz et al., 2018). In porcine kidney AOC1 the de-glycosylation had no effect on enzyme activity and protein stability (Schwelberger and Bodner, 1999). Similar to its homologous hAOC1^{Asn110} site, the hAOC3^{Asn137} site is mainly occupied by oligomannosidic glycans (Gludovacz et al., 2018). Mutational studies were done to prevent glycosylation in hAOC3, and they showed that hAOC3^{Asn592}, hAOC3^{Asn618} and hAOC3^{Asn666} regulate the enzymatic activity of hAOC3 (Maula et al., 2005). Moreover, is known that sialyated glycans are involved in the interaction between hAOC3 and leucocytes (Maula et al., 2005; Salmi and Jalkanen, 1996).

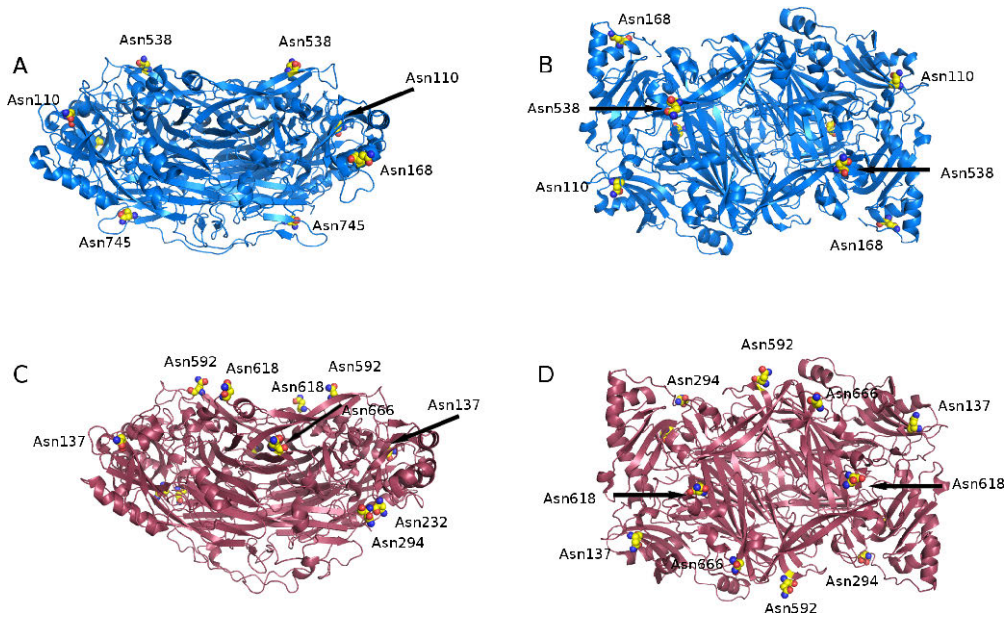


Figure 7. N-glycans distribution on hAOC1 and hAOC3. The glycosylated Asn (yellow spheres) and are distributed throughout the surfaces. A: side view and B: top view of hAOC1 (PDB ID 3HI7) (McGrath et al., 2009). C: side view and D: top view of hAOC3 (PDB ID 4BTX) (Bligt-Lindén et al., 2013).

2.7. Human CAOs as biomarkers and in drug development

2.7.1. AOC1

AOC1 is implicated in histamine intolerance due to its role in the metabolism of ingested histamine (Maintz and Novak, 2007). The AOC1 activity increases significantly in pregnancy. The exact role is still unclear, but it has been suggested to be a metabolic barrier that prevents the entry of histamine from the placenta into the maternal or fetal circulation (Maintz et al., 2008). Low AOC1 activity during pregnancy it is linked to high risk pregnancies (Maintz et al., 2008). The intra-individual variations of AOC1 levels during pregnancy suggest that the levels of AOC1 might not have prognostic significance if measured during pregnancy (Maintz et al., 2008). Due to the high costs of an AOC1 assay, the measurement of the AOC1 levels is not yet standard procedure during pregnancies. In 2016, an

assay that accurately quantifies AOC1 in various fluids was developed and this assay might provide a stepping stone to explore the use of AOC1 as a biomarker (Boehm et al., 2016). For example, it has been suggested that the levels of AOC1 in non-pregnant women can be used to indicate women at risk for complicated pregnancies (Maintz et al., 2008).

Low AOC1 activity might also act as a possible indicator for intestinal mucosa damage in inflammatory and oncological diseases (Schmidt et al., 1990; Raithel et al., 1998). Additionally, AOC1 is an off-target of different commonly used drugs, such as the antidepressant amitryptyline and the antibiotic clavulanic acid, and thus the intake of these drugs should be taken into account in the interpretation of histamine intolerance symptoms and AOC1 concentrations (Maintz and Novak, 2007). In the assay developed by Boehm et al. (Boehm et al., 2016) that quantifies hAOC1 in different biological fluids, the effect of the off-target inhibition of AOC1 was tackled. Since the potent inhibitors diminazene aceturate and aminoguanidine do not interfere with the measurements. Altogether, AOC1 has the potential to be used as a biomarker.

2.7.2. AOC3

hAOC3 is linked with several diseases and it is a promising drug target (Boomsma et al., 2003; Dunkel et al., 2011) with the potential use against inflammatory conditions but also against tumour progression and the metastatic spread of cancer (Jalkanen and Salmi, 2017; Salmi and Jalkanen, 2011). The role of hAOC3 in inflammatory diseases, diabetes and cancer has been extensively studied using animal models (reviewed in Jalkanen and Salmi, 2017). AOC3 can be inhibited by small molecule inhibitors (Foot et al., 2012; Inoue et al., 2013b; Jarnicki et al., 2016; O'Rourke et al., 2008; Schilter et al., 2015; Yamaki et al., 2017c, 2017b, 2017a) as well as by antibodies (Arndtz et al., 2017; Vainio et al., 2005). However, due to the species-specific differences in the active site, *in vitro* testing and testing in animal models can become problematic. For example, the pyridazinone inhibitors are potent inhibitors of hAOC3 but weak inhibitors of mouse AOC3 (Bligt-Lindén et al., 2013).

Soluble AOC3 (sAOC3) has the potential clinical use as a biomarker (Pannecoeck et al., 2015), despite that it is not as a general inflammation marker (Aalto et al., 2012). Elevated levels of sAOC3 can be used as a marker for cardiovascular calcifications (Aalto et al., 2012), liver, gastric and colorectal cancers (Kemik et al., 2010; Li et al., 2011; Yasuda et al., 2011) and to predict cancer mortality in colorectal cancer in type-2 diabetes patients (Li et al., 2014). Another application for AOC3 as biomarker takes advantages of its adhesion function (Aalto et al., 2011). AOC3 is stored inside of the cells but upon inflammatory stimuli it is transferred to the endothelial cell surface where it prevails during inflammation (Salmi and Jalkanen, 2014). When the hAOC3 leucocyte ligands, Siglec-10 (Kivi et al., 2009) and Siglec-9 (Aalto et al., 2011), were found the use of AOC3 as a target for visualising inflammation using PET imaging was developed. A Siglec-9 derived peptide that fits in the enzymatic groove of hAOC3 (Aalto et al., 2011) was used in a variety studies (Aalto et al., 2011; Jaakkola et al., 2000; Jensen et al., 2017; Li et al., 2013; Virtanen et al., 2015).

3. Aims

This thesis mainly concentrates on the characterization of mammalian CAOs but also cyanobacterial PAOs were analysed. The specific aims for this thesis were as it follows:

I. To study the evolutionary relationship of mammalian CAOs in order to aid in the correct annotation of new and poorly annotated CAOs and to derive rules to classify AOC1-4 and to reveal residues that are important for the substrate preference (publication I).

II. To study the structure and function of the hAOC-1 attached N-glycans with a particular emphasis on the conserved N-glycosylation site with high-mannose glycosylation in hAOC1 and hAOC3 (publication II).

III. To predict and analyse the interactions between hAOC3 and the Siglec-9 derived peptide (publication III) to suggest improvements of hAOC3-targeted biomarkers.

IV. To study the substrate preference and catalytic reaction mechanism of the *Synechocystis sp.* PCC 6803 PAO (SynPAO) and give structural explanation for the functional characteristics using phylogenetic study, homology modelling and structural analysis (publication VI).

4. Materials and methods

4.1. Protein sequence and structure databases

All the amino acid sequences and protein structures used in this thesis were collected from web-based databases. The amino acid sequences were collected from UniProtKB ("UniProt: the universal protein knowledgebase," 2017) and from the NCBI Protein database (https://www.ncbi.nlm.nih.gov), whereas the protein structures were retrieved from the PDB (www.rcsb.org) (Berman et al., 2000).

Sequence searches were done using the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) at NCBI, against non-redundant protein sequences database. The BLAST search finds regions of local similarity between sequences and calculates the statistical significance of the match.

4.2. Sequence alignments

All the alignments in this thesis were made in the BODIL modelling environment (Lehtonen et al., 2004). In publications I and IV, VERTAA (Johnson and Lehtonen, 2000) was used to generate a structure-based sequence alignment, which worked as a starting point for the multiple sequence alignments. The usage of a pre-aligned structure-based sequence alignment increases the reliability of the residue-residue correspondence and the correct positioning of gaps in the final multiple sequence alignment. The multiple sequence alignments were done using MALIGN (publication I, III and IV) (Johnson et al., 1996). The pairwise alignment used for the homology modelling of porcine AOC1 (publication II), was done using MALIGN (Johnson et al., 1996).

4.3. Sequence and phylogenetic analysis

The amino acid sequences were submitted to SignalP (publication I and III) (Petersen et al., 2011) for signal peptide prediction and to TMHMM (publication I) (Sonnhammer et al., 1998) and SOSUI

(publication IV) (Hirokawa et al., 1998) for transmembrane helix prediction. Skylign (publication I) (Wheeler et al., 2014) was used to create the sequence logos representing the profile hidden Markov models of each alignment. For the phylogenetic study in publication I, sequences were curated manually. Sequences without the catalytic aspartate, the tyrosine that is post-transcriptionally modified to TPQ, and the tree histidines that coordinate the copper ion were removed. This ensured that the sequences analysed are catalytically active. Additionally, isoforms with the same active site motif were removed as they induce redundancy in the trees.

Phylogenetic trees were constructed using the Maximum Likelihood (ML) (publications I and V) and Neighbour Joining (NJ) (Saitou and Nei, 1987) (publication I). Due to the size of the dataset in publication I, MEGA-CC (Kumar et al., 2012) was used as it allows parallel execution on multiple processors and cores, whereas MEGA6 (Tamura et al., 2013) was used in publication V. The NJ trees were generated using the Jones-Taylor-Thornton (JTT) substitution matrix (Jones et al., 1992). For the ML trees, the best-fit evolutionary model was assessed using the Bayesian Information Criteria (BIC). To generate the trees in publication I, different models were chosen based on the BIC score: 1) Poisson correction model (Zuckerkandl and Pauling, 1965) with gamma distribution and invariant sites (P+G+I) for the tree containing all CAOs, 2) JTT with gamma distribution (JTT+G) for the trees with AOC1 and AOC2 proteins and 3) JTT with gamma distributions and invariant sites (JTT+G+I) for the tree with AOC3 and AOC4 proteins. For publication IV, the Lee and Gascuel (Le and Gascuel, 2008) model with discrete gamma distributions and invariant sites (LG+G+I) was used. The stability of the topology of the phylogenetic trees was evaluated using the Felsenstein's bootstrap replication method (Felsenstein, 1985) with 500 replications (publication I and IV).

To aid in the classification of CAOs and to complement the phylogenetic analysis, multivariate analyses was performed by the co-author Prof. Mark S. Johnson (publication I). The distance data was calculated using the JTT substitution matrix (Jones et al., 1992) and supplied to a program to perform a multivariate analysis (PCA,

MS Johnson). The program returns a pseudo-PDB coordinate file which was visualized using BODIL (Lehtonen et al., 2004).

4.4. Modelling

All homology models used in this thesis were created with MODELLER (Sali and Blundell, 1993). In publication I, the hAOC2 model was generated based on the crystal structure of hAOC3 (PDB ID 4BTX) (Bligt-Lindén et al., 2013) and, in publication II, the model of porcine AOC1 (pAOC1) was generated based on the crystal structure of hAOC1 (PDB ID 3HI7) (McGrath et al., 2009). Of the 10 generated models for each protein, the model with the lowest MODELLER objective function was chosen for further analysis (publication I and II). Visual inspection of the homology models was done by superimposing the model and respective template using VERTAA in BODIL (Johnson and Lehtonen, 2000) (publication I and II). The models were evaluated using PROCHECK (publication I and II) (Laskowski et al., 1993), QMEAN (publication I) (Benkert et al., 2009), Verify3D (publication II) (Wallner and Elofsson, 2003) and ProSA-web (publication II) (Wiederstein and Sippl, 2007).

The model of hAOC2 (publication I) has 90% of residues in the favoured regions of the Ramachandran plot, QMEAN score of 0.704 (acceptable values range 0-1, the higher value the better), the model and template have a root mean square deviation (RMSD) of 0.67 Å for the C-alpha atoms of the superimposed structures and share a sequence identity of 68%. For the model of pAOC1 (publication II), 90.7% of residues are in the most favoured region of the Ramachandran plot. Based on the Verify3D analysis 85.1% of the residues have an average 3D to 1D score of 0.2 or more (threshold to pass the test is at least 80% or amino acids with score of 0.2 or more) and ProSA-web gave a Z-score of 9.88, which in the range of scores typically found for native proteins of similar size. Thus, the models were considered reliable for the 3D structural analysis of hAOC2 and pAOC1.

The solvent accessibility of the N-glycosylated Asn residues was determined by using NACCESS (publication II)

(http://bioinf.manchester.ac.uk/naccess/) using a probe size of 5 Å to simulate the accessibility of the glycosylating enzymes into the glycosylation sites.

4.5. Ligand docking

In publication III, the Siglec-9 peptide used as ligand was generated using BIOVIA Discovery Studio v4.5 (San Diego, CA, www.accelrys.com). The collaborators, Dr. Jerôme de Ruyck and Prof. Gerard Vergoten from the University of Lille did the molecular docking simulation and the post-docking energy minimization (see publication III materials and methods). The best ranking 20 poses were evaluated by us. We analysed the docking poses based on previously published work on Siglec-9 peptide (Aalto et al., 2011), extracellular part of Siglec-9 (Elovaara et al., 2016) and comparison with hAOC3 crystal structures (PDB IDs 2Y73, 2Y74 and 4BTY) (Bligt-Lindén et al., 2013; Elovaara et al., 2011b). This limits the possible poses to those where R3 and R9 interact with the active site cavity, R3 binds near TPQ and the terminal Lys is located near the protein surface. The interactions were predicted using the PLIP server (Salentin et al., 2015) and Maestro (Schrödinger, LLC).

In publication IV, the spermine structure was downloaded from the PubChem database (Kim et al., 2016) and prepared using the LigPrep module in Maestro with default settings (LigPrep, version 3.8; Schrödinger, LLC). The spermine molecule was removed from the homology model of SynPAO and the model was prepared according to the protein preparation protocol in Maestro (Maestro, version 10-6; Schrödinger, LLC). A docking grid with a positional constraint (4.0 Å to the hydroxyl group of Tyr403) was generated to cover the active site volume. The docking of spermine to the active site of SynPAO was done using Glide (Friesner et al., 2004), ten docking poses were generated and ranked based on Emodel score. All poses were visually analysed and compared with the spermine complex crystal structures of ZmPAO (PDB ID 3KU9; (Fiorillo et al., 2011)) and yeast Fms1 (PDB ID 1XPQ; (Huang et al., 2005)), and they were also compared with the homology model of mouse SMO with docked spermine (Tavladoraki et al., 2011).

4.6. Visualization

PyMOL (DeLano, 2002) was used for visualizing the 3D structures and models (publication I, II, III and IV) and to generate high-resolution figures. The labels were added with GNU Image Manipulation Program (version 2.8).

4.7. Experimental work

My co-authors performed the experimental work for the publications included in this thesis. The experiments are explained in detail in the Materials and Methods section of the original publication (publication II, III and IV).

5. Results and discussion

5.1. Mammalian copper amine oxidases

The role of human CAOs in various diseases, their 3D structure and substrate preferences has been extensively researched (Finney et al., 2014), in particular, hAOC1 and hAOC3. Despite the known differences in the substrate preferences, CAOs are currently divided into two sub-families E.C. 1.4.3.21 (primary amine oxidases) and E.C. 1.4.3.22 (diamine oxidase). Membrane bound CAOs with activity towards monoamines are commonly referred to as semicarbazide sensitive amine oxidases (SSAO), which is a poor convention as all CAOs are inhibited by semicarbazide (Elmore et al., 2002). The soluble plasma CAOs with activity towards monoamines are also referred to as soluble amine oxidase (SAO)/benzylamine oxidases (BAO). This thesis aims to address the ambiguities in the current naming system by creating a more specific classification system for CAOs. Additionally, it aims to understand which residues in the active site contribute to the substrate preference of each CAO subfamily.

5.2. Evolution of vertebrate CAOs

The first extensive phylogenetic analysis of vertebrate CAOs was made in publication I. The dataset consists of 369 vertebrate CAOs sequences (publication I) with the sequence identity ranging between ~30% and 99%. To infer the trees two methods were used: ML and NJ. As both trees were congruent, the ML tree was used for further analysis (Figure 8A). The phylogenetic tree shows that vertebrate CAOs are divided into two major branches: (1) AOC1 proteins and (2) AOC2-4 proteins. The tree shows that AOC2 and AOC3/4 proteins are more closely related to each other than to AOC1 (Figure 8A).

AOC1 proteins were found in all bony vertebrates. Based on the phylogenetic tree, all AOC1s share the same clade, supported with bootstrap 100% (Figure 8A). This suggests a similar substrate

preference for the fish, amphibian, reptile, bird and mammalian AOC1 protein. Almeida and Beaven (1981) reported that AOC1 has activity towards histamine in the brain of lower vertebrates, which supports the tree. AOC2-4 proteins were found in Amphibians and Amniotes, whereas AOC4 was found only in placental mammals. Unlike AOC1, the AOC2 and AOC3 proteins were distributed in different clades (Figure 8A). The mammalian and non-mammalian proteins were in different sub-clades, supported with bootstrap 85% and 99% for the non-mammalian and mammalian proteins, respectively (Figure 8A).

CAOs from *Rhincodon typus* (whale shark) and *Callorhinchus milii* (elephant shark), which are cartilaginous fishes, were used as out-group for the phylogenetic analysis (publication I). *C. milii* has diverged ~450 million years ago and it is the slowest evolving genome of all known vertebrates, making it a good model to study the ancestral state of genomes (Venkatesh et al., 2014). *C. milii* CAO is annotated as "retina-specific amine oxidase" and this protein has the highest sequence identity with hAOC2 (~44%), followed by hAOC3 (~43%) and hAOC1 (~41%). Similarly, the sequence identity of *R. typus* is also the highest with hAOC2 (~43%) followed by hAOC3 (~42%) and hAOC1 (~39%).

The previous phylogenetic studies published on CAOs either used a small number of sequences (ten CAO sequences) (Lai et al., 2012) or were focused on the porcine genome (Schwelberger, 2010, 2006). Thus results of these studies are very limited. The results from publication I, provide a more in-depth knowledge of the phylogenetic relationships between mammalian and non-mammalian CAOs. Since a larger number of CAO sequences from a variety of vertebrate species were used in this study.

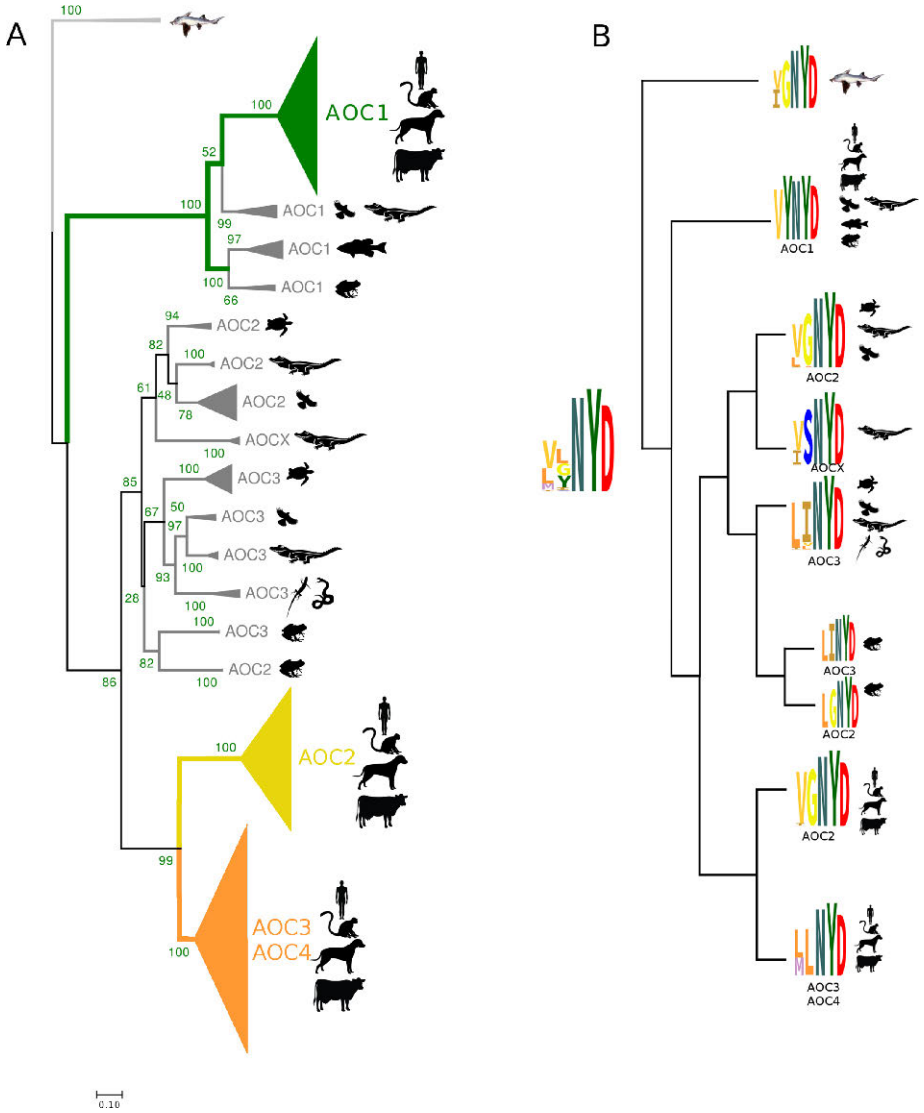


Figure 8. A: ML phylogenetic tree of vertebrate CAOs. B: ML tree topology with the corresponding HMM motif logos (Figure from publication I).

5.3. Motif-based classification of mammalian CAOs

The first structural studies involving CAOs were made using non-mammalian sources. In the studied species residue X2 of the active site motif (X1-X2-Asn-Tyr-Asp) is always a small hydrophobic residue: Gly in ECAO, *Hansenula polymorpha* CAO, AGAO and *Pichia pastoris* CAO, and Ala in *Pisum sativum* CAO. The structural work with hAOC3 showed that X2 is a Leu, which acts as a "guardian" residue (Airenne et al., 2005; Salminen et al., 1998). In fact, in the mammalian proteins, residue X2 seems to be variable in the different CAOs, and it seems to follow the different substrate preference of mammalian CAOs: a Tyr in AOC1, a Gly in AOC2, and a Leu in AOC3 (Airenne et al., 2005; Kaitaniemi et al., 2009; Salminen et al., 1998). Altogether, this motivated to study if the residues X1 and X2 could be used to distinguish between the mammalian CAOs (publication I). Since the phylogenetic tree shows that the conservations of the motif in vertebrates is in accordance with the substrate preference (Figure 8A), residues X1 and X2 were analysed with more detail, with special emphasis on mammalian CAOs.

The active site motif was analysed using the HMM motif logos, and related to the tree topology (Figure 8B). Results show that in mammals X2 is always a Tyr in AOC1, a Gly in AOC2 and a Leu in AOC3/ACO4. Similarly, in the remaining vertebrates residue X2 is a Tyr in AOC1 and the majority of AOC2s have a Gly. In AOC3s, X2 is less conserved (Figure 8B). It is always hydrophobic but is more frequently Ile than Leu in non-mammalians, and some species have a Val. Additionally, a different motif was found in some crocodilian CAOs, where X2 is a Ser. These proteins were identified as AOCX (Figure 8B) and likely have different substrate preference.

Mammals are the only species with AOC3 and AOC4 (Figure 8A), which both have a Leu in position X2 (Figure 8B). Thus, the classification of these proteins was further studied. The multiple sequence alignment containing only mammalian CAOs showed that the average sequence identity of mammalian AOC1s to AOC2s is 62% and to AOC3/AOC4 it is 60%, whereas the identity of AOC2 to AOC3/AOC4 is 73%. Regarding the AOC3 and AOC4 proteins, the

sequence identities range between 84% and 98%, which shows that these proteins are highly similar. The multivariate plot with mammalian CAOs showed three distinct clusters, one per CAO subfamily, which is in accordance with the phylogenetic tree of vertebrate CAOs (Figure 9A). Thus, it did not help in the classification of AOC3 and AOC4. The residue at position X1 in the motif was analysed next. When comparing the sequences of hAOC3 and bAOC4 with the gene information on the collected AOC3 and AOC4 proteins, it was found that proteins with a Leu as X1 mostly derive from a different gene than proteins with a Met as X1. Therefore, proteins with a Leu as X1 were marked as AOC3 and proteins with a Met as AOC4 and a multivariate plot was generated. In the resulting plot, proteins with a Leu cluster separately from the proteins with a Met (Figure 9B) suggesting that CAOs with a Leu should be classified as AOC3 and with a Met as AOC4.

The generated HMM motif logos for each mammalian CAO subfamily show the conservation of motif within the same subfamily (Figure 9C). Altogether, results from publication I show that the TPQ signature motif can indeed be used for the classification of mammalian CAOs. This motif-based classification is consistent with the different substrate preferences. The Tyr at position X2 in AOC1 is consistent with the diamine preference of AOC1s, the Gly at position X2 in AOC2 is consistent the substrate preference for aromatic monoamines, the Leu at X2 at AOC3/AOC4 with the aliphatic monoamine preference, and, furthermore the Leu and Met at X1 differentiate between the small aliphatic amines and aliphatic polyamine preference in AOC3 and AOC4, respectively.

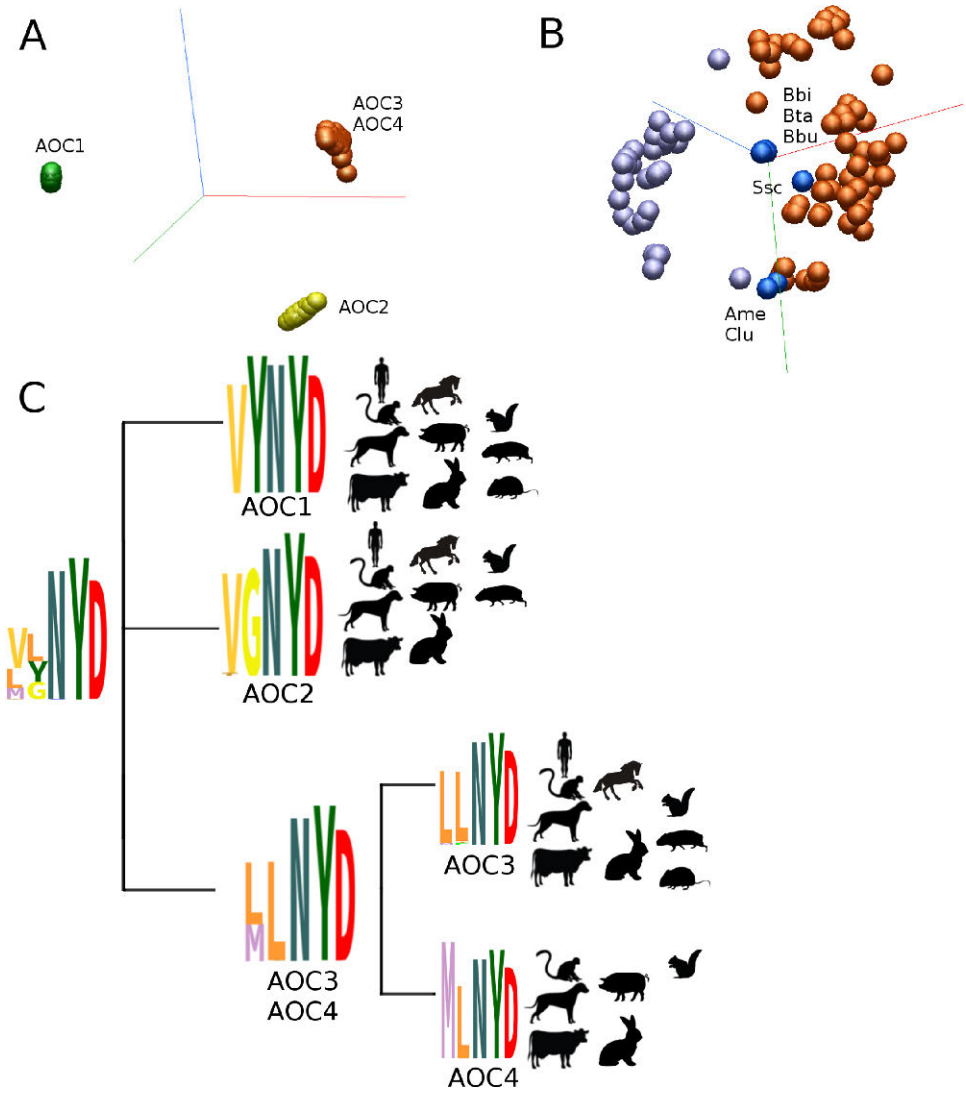


Figure 9. A: multivariate plot of mammalian CAOs showing the different subfamilies. B: multivariate plot of mammalian AOC3 and AOC4 proteins, blue sphere represent outliers. C: Tree topology based on Figure 8A, with the HMM motif logo of mammalian CAO proteins. Bbi – *Bison bison bison*; Bta- *Bos taurus*; Bbu- *Bubalus bubalis*; Ssc – *Sus scrofa*; Ame- *Ailuropoda melanoleuca*; Clu – *Canis lupus familiaris* (Figure from publication I).

5.4. Phylogenetic analysis of CAO subfamilies

In the light of the motif-based classification, the CAO subfamilies could be analysed in more detail. This included a phylogenetic analysis for each subfamily. The trees were generated separately for the AOC1, AOC2 and AOC3/4 subfamilies. In the generated trees, the branches are mostly supported by low bootstrap values (Figure 10-12). This generally occurs when the time frame is short, as there are not enough substitutions supporting each node and, thus, the dataset as a whole remains sensitive to random fluctuations in the observed mutations and small differences (Soltis and Soltis, 2003). All the trees have a different topology, which can be seen in rodent CAOs of the AOC1 and AOC2 trees. In AOC1, the rodents are divided into the two main branches of the tree (Figure 10). Myomorpha (mouse-like rodents) rodents form their own branch, whereas rodents from the Hystricomorpha, Sciuromorpha and Castorimorpha suborders share the same branch as primates and *Oryctolagus cuniculus* (rabbit). *Rattus norvegicus* (rat; Myomorpha) and *Cavia porcellus* (guinea pig; Hystricomorpha), which both are rodents but in different branches of the tree are affected differently by the tricyclic antidepressant amitriptyline (Rajtar and Irman-Florjanc, 2007). Amitriptyline acts as an inhibitor of *R. norvegicus* AOC1, whereas it increased the activity of *C. porcellus* AOC1 leading to a faster histamine inactivation. Their ligand binding properties reflect the branching of the phylogenetic tree for AOC1 proteins (Figure 10). On the other hand, all rodents AOC2s group together and share the same branch as *O. cuniculus* (Figure 10), which shares the same sub-branch as primates.

Due to the high average sequence identity between AOC3 and AOC4 (81%) and the high conservation of the D4 catalytic domain (59-100%), the phylogenetic tree for these proteins was generated together (Figure 12). Based on the tree it seems that the divergence of AOC3-AOC4 proteins occurred in a species-specific way, likely due to different evolutional pressure. For example, AOC3 and AOC4 from monkeys do not share the same most recent last common ancestor, whereas bovine AOC3 and AOC4 share (Figure 12). It has been suggested previously that species with the *AOC4* gene have

high level of serum CAO (Schwelberger, 2006). Among these species are *Ovis aires* (sheep), *Bos taurus* (cow), *O. cuniculus*, *Capra hircus* (goat), *Canis lupus familiaris* (dog) and *Sus scrofa* (pig) (Boomsma et al., 2003; Schwelberger, 2006). In general, this is in accordance with the results in publication I. However, only one gene was found in *C. lupus familiaris* and *S. scrofa*. More details about the outliers in this study can be found in publication I. The origin of AOC4 has been traced to duplication of the *AOC3* gene (Schwelberger, 2010). In accordance, no AOC4 with the AOC3 motif (X1=Leu) were found, but AOC3s were found with a Met as X1. The AOC4 proteins from *O. cuniculus* and *S. scrofa* have high activity towards mescaline, a hallucinogenic plant-derived amine, suggesting a dietary influence on the metabolized substrates (Lyles, 1996). Altogether, results from publication I suggest that CAOs with predicted AOC4 activity (X1=Met), but deriving from the *AOC3* gene should not be named as AOC3 but as AOC4-like. Furthermore, it is of high importance to analyse the genomic origin of the encoded protein.

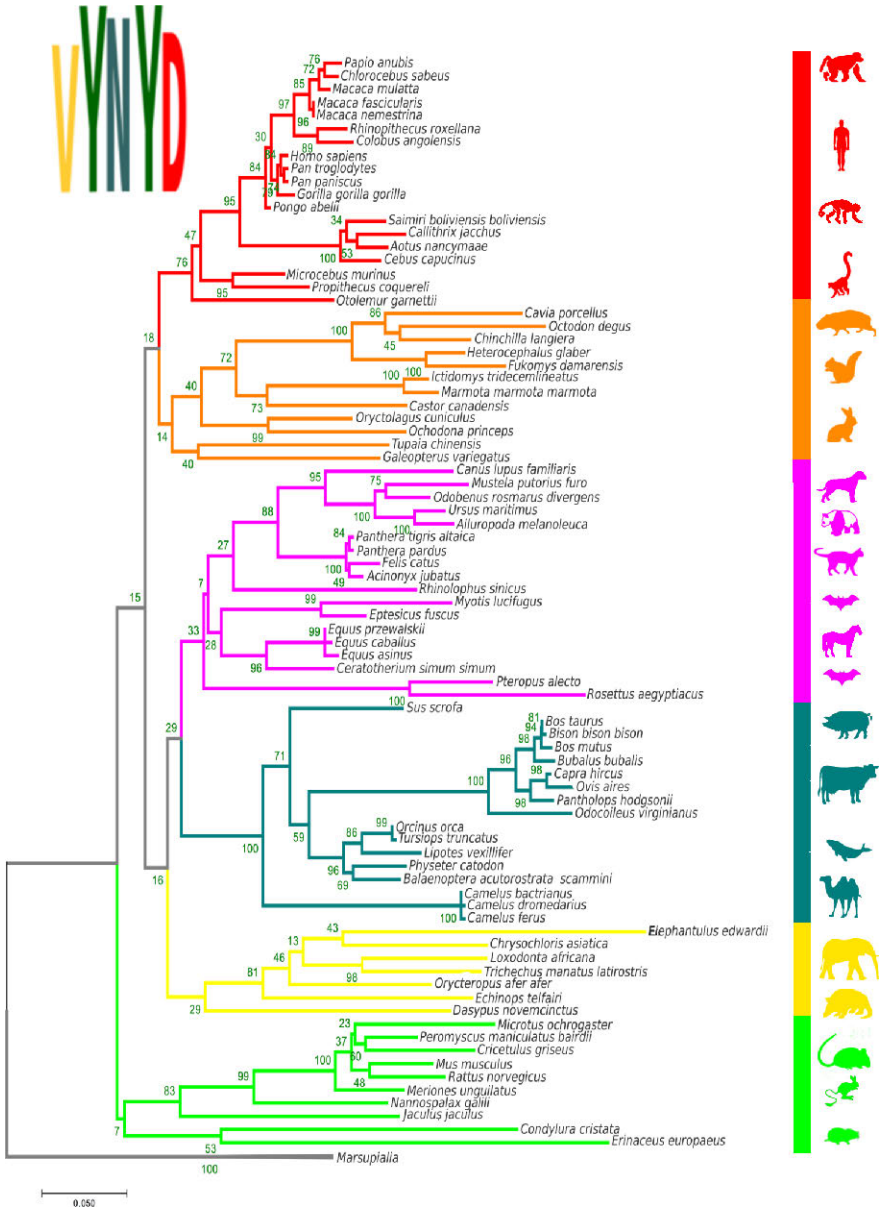


Figure 10. Phylogenetic tree of mammalian AOC1 proteins with HMM motif logo in the upper left corner (Figure from publication I).

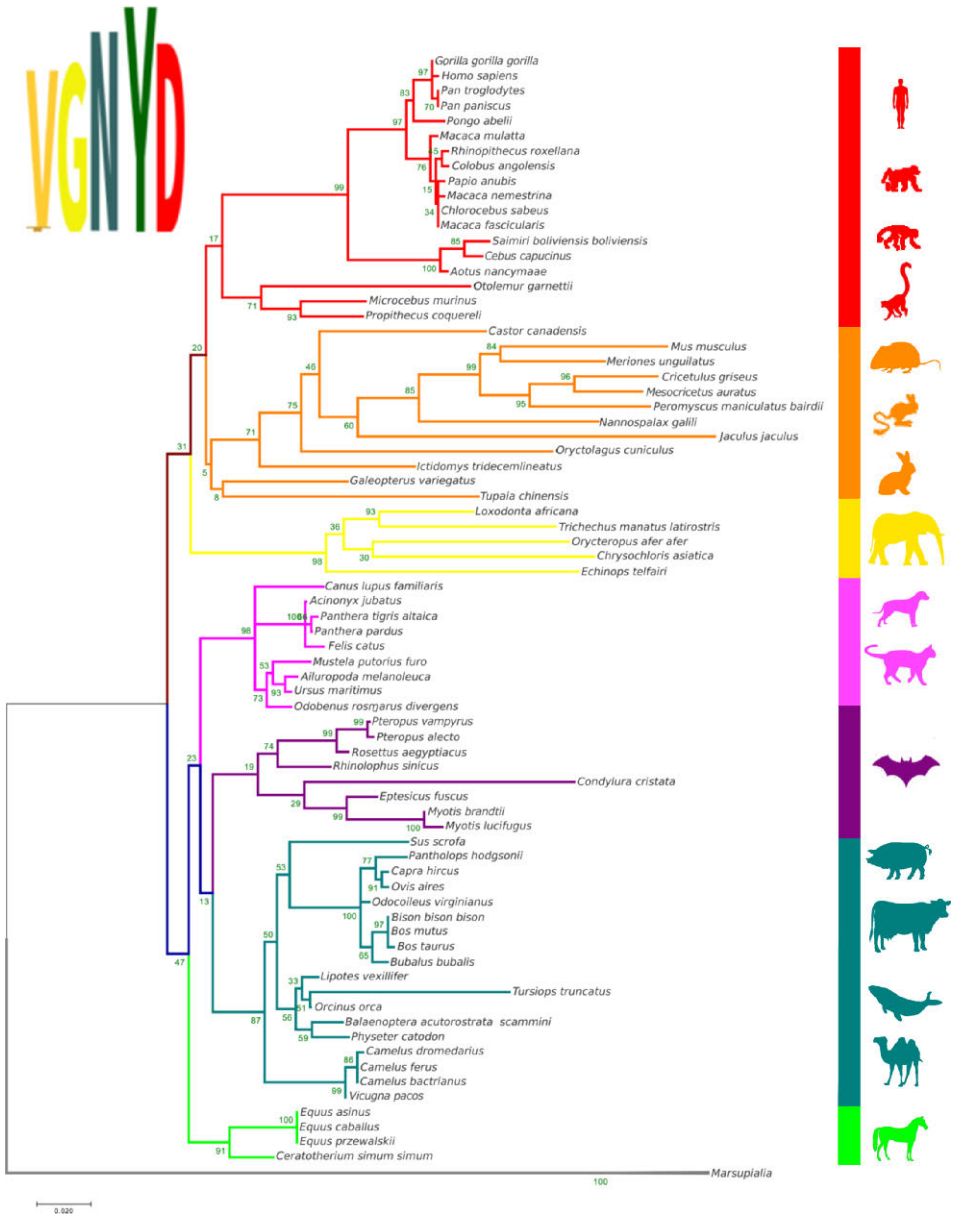


Figure 11. Phylogenetic tree of mammalian AOC2 proteins with HMM motif logo in the upper left corner (Figure from publication I).

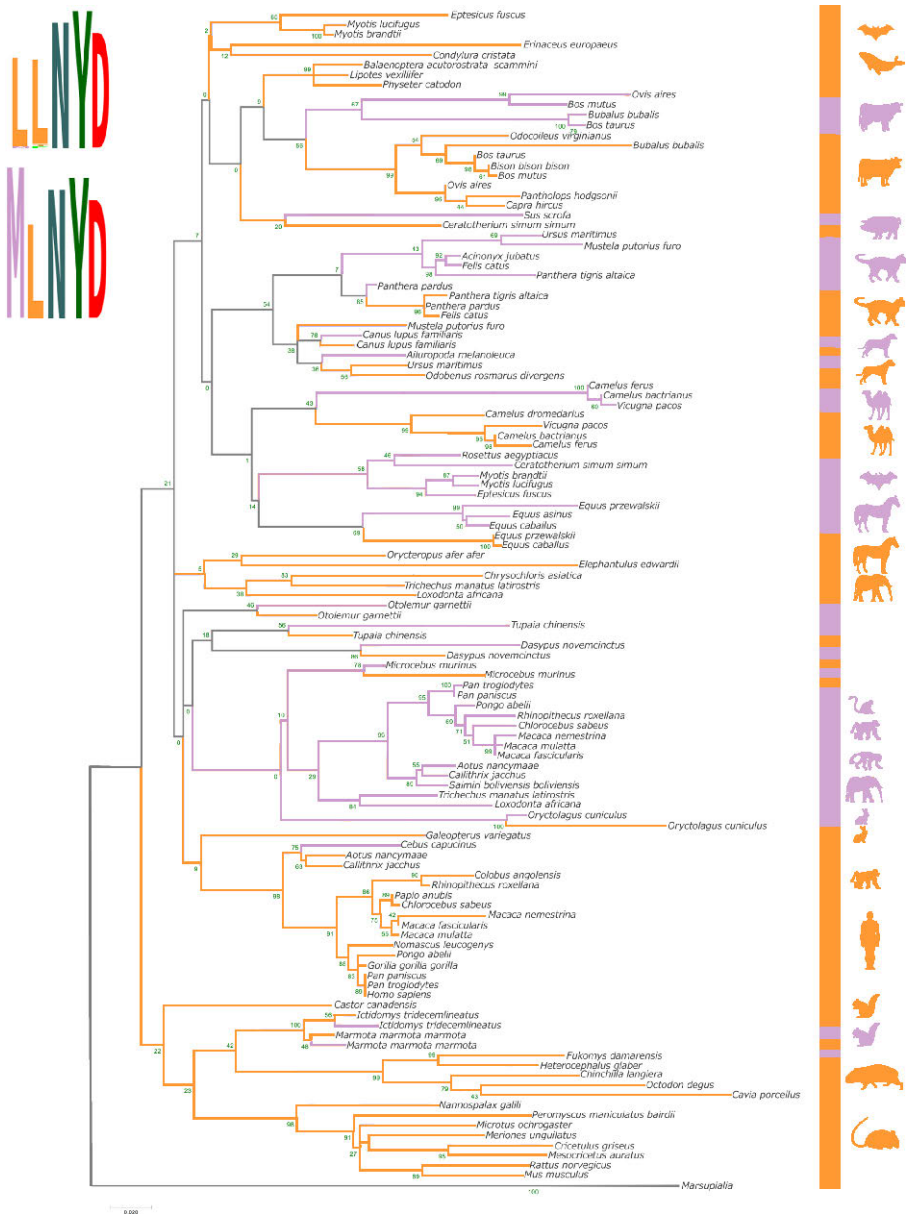


Figure 12. Phylogenetic tree of mammalian AOC3 (orange) and AOC4 (violet) proteins with HMM motif logos in the upper left corner (Figure from publication I).

5.5. Active site of mammalian CAOs

The area surrounding TPQ was studied in the 3D structures of hAOC1 (PDB ID 3MPH; McGrath et al., 2009), hAOC3 (PDB ID 4BTX; Bligt-Lindén et al., 2013) and bAOC4 (PDB ID 1TU5; Lunelli et al., 2005), and in the homology model of hAOC2. Half of the residues near TPQ are fully conserved (Figure 13), and most of them are involved in TPQ biogenesis and/or catalytic mechanism (Airenne et al., 2005; Ernberg et al., 2010; Klema and Wilmot, 2012). The catalytic Asp (Aps373 in hAOC1) and a nearby Tyr (Tyr371 in hAOC1) make the conserved Y-X-D motif. The Tyr in this motif has been identified as a "gate" residue in mammalian CAOs (Kaitaniemi et al., 2009; Salminen et al., 1998). This Tyr likely interacts with the amine substrates of mammalian CAOs similarly, as it interacts with the pyridine ring in the crystal structure of hAOC3 with 2HP (Jakobsson et al., 2005) and it possibly interacts with the aromatic ring of benzylamine in hAOC3 (Kaitaniemi et al., 2009). TPQ is oriented by two conserved Tyr (Tyr359 and Tyr463 in hAOC1) and a Asn (Asn460 in hAOC1) (Airenne et al., 2005; Ernberg et al., 2010), whereas the conserved Ser in the vicinity of TPQ (Ser465 in hAOC1) forms hydrogen bonds that stabilize the active site architecture (Figure 13A). The conserved hydroxyl containing residue (Ser/Thr; Thr457 in hAOC1) and the Asp from the motif (Asp462 in hAOC1) form an hydrogen bond. This interaction is part of an inter-monomeric hydrogen bond network that together with fully conserved Arg (Arg492 in hAOC1) and His (His430 in hAOC1) from the other monomer stabilizes the tip of the β-hairpin. Similar hydrogen bonding network has been identified also in ECAO (Parsons et al., 1995).

The non-conserved residues in the area surrounding TPQ are likely responsible for the different substrate preference of CAOs (Figure 13). The diamine preference is unique to hAOC1 (Elmore et al., 2002), and based on the phylogenetic tree (Figure 8), the AOC1 subfamily shares the substrate preference. AOC1 proteins have a set of uniquely conserved residues that seem to contribute to the diamine preference (Val458, Tyr459 and Asp186 in hAOC1). Tyr459 in hAOC1 together with a conserved aromatic residue (Trp367 in

hAOC1) are in a suitable position to orient the substrates, whereas Asp186 is suitable for binding the second diamine group. This is in accordance with a study made by McGrath and co-workers (2009), which predicted Asp186 to be involved in histamine binding. Val458 is also conserved in AOC2s, which suggests that it contributes to the aromatic substrate preference. The extra space in the active site of AOC2, due to the small residues Gly (X2) and Val (X1) (Gly463 and Val462 in hAOC2) from the active site motif, provides an optimal environment for the aromatic ring of the substrates. In fact, Gly has been shown to be important for p-tyramine binding, since after the mutation of Leu in hAOC3 to a Gly, hAOC3 acquired activity towards p-tyramine, which the hAOC3 wild type did not have (Elovaara et al., 2011). In the same position as Asp186 in hAOC1, there is a His in hAOC2 (His206). This residue could interact with polar and/or aromatic substrates. The His is conserved in AOC2s, with the exception of bats that have a Gln. Thus, it can be postulated that His is specific to AOC2 and contributes to its substrate preference.

AOC3 prefers small aliphatic amines, whereas AOC4 long polyamines (Bonaiuto et al., 2010; Elovaara et al., 2011; Kaitaniemi et al., 2009). In both proteins, X2 is a Leu, which has been associated with blocking/allowing access to TPQ (Airenne et al., 2005; Jakobsson et al., 2005; Kaitaniemi et al., 2009; McGrath et al., 2009). Mutational studies of this residue have shown that it is involved in the substrate preference of hAOC3 (Elovaara et al., 2011; Kaitaniemi et al., 2009). This Leu together with a conserved Phe in AOC3 and AOC4 (Phe389 in hAOC3), provide a hydrophobic environment in the active site that can accommodate the carbon tails of aliphatic amines (Figure 13C and D). The residue X1 is a Leu in AOC3 and a Met in AOC4. The exact role of this residue is not yet known, but the difference provides the active site of AOC4 a slightly more polar nature, when compared to AOC3 (Figure 10). In the same position as Asp186 in hAOC1, there is a conserved Thr in AOC3 (Thr212 in hAOC3), and an almost conserved Asn in AOC4 (Asn211 in bAOC4) (Figure 13C and D). Moreover, in hAOC3, it has been shown by mutational studies that Thr212 merely has a role as an additional

gate, contributing only slightly to substrate preference (Elovaara et al., 2011).

Rodents have been used as model organism in human AOC3 related research, for diseases, drug design and *in vivo* imaging, whereas AOC4 related research has been focusing on bAOC3. The studies on the substrate preference of CAOs have been mainly done using human and rodent proteins (Elmore et al., 2002; Finney et al., 2014; Lyles, 1996; Shen et al., 2012; Zhang et al., 2003). The species-specific differences that have been noticed in the drug design projects are known the be problematic (Bligt-Lindén et al., 2013; Foot et al., 2012; Inoue et al., 2013a, 2013b). The comparison of rodent and primate proteins suggests that for small substrates and inhibitors binding covalently to TPQ, rodents and other primates might serve as a good model for humans, as the active site in that region is quite conserved. However, for larger inhibitors, the active site channel of the model organism should be studied in detail. Moreover, besides the species-specific properties of the active site, it is crucial to take into account the other CAO proteins expressed in the model organism.

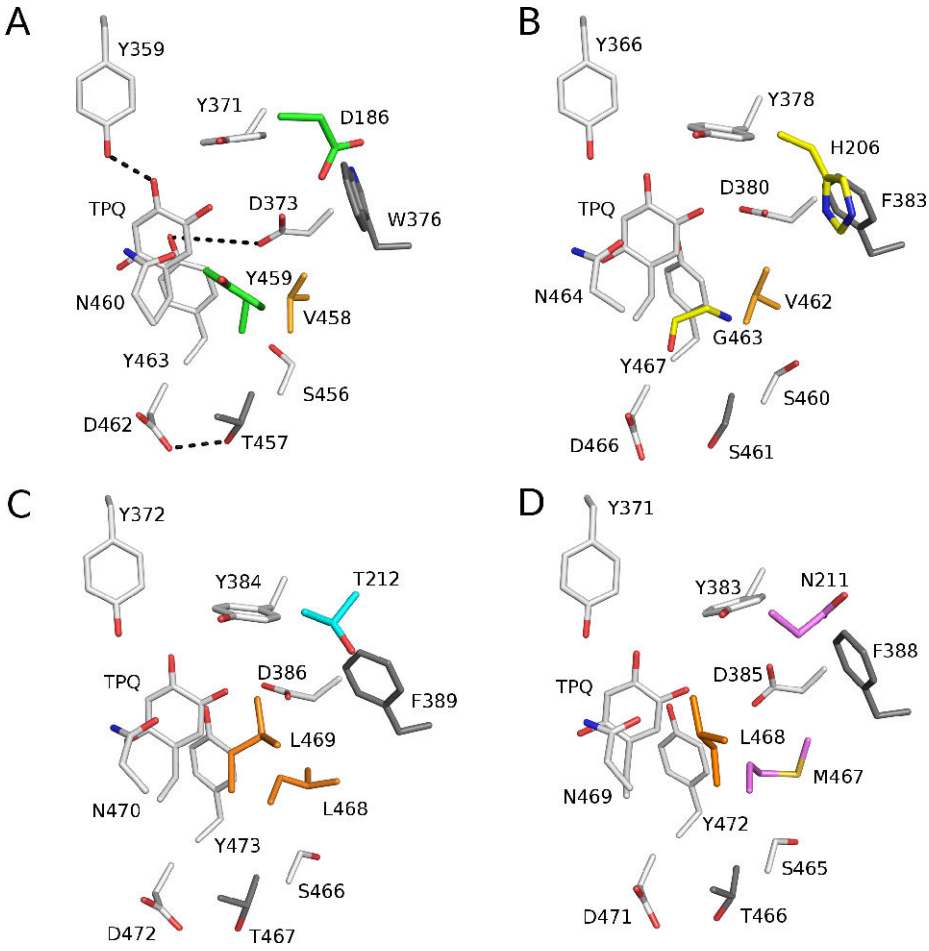


Figure 13. Active site of the different mammalian CAOs. A: hAOC1; B: hAOC2; C: hAOC3; D: bAOC4. (Figure from publication I).

5.6. The conserved N-glycosylation site in CAOs

A N-glycosylation site is among the conserved characteristics of mammalian CAOs. The site is located in the β1.3 β-strand that precedes the loop linking the two β-strands in the central four-stranded β-sheet of the D2 domain. This glycosylation site, which is hAOC1^{Asn110}, pAOC1^{Asn115} and hAOC3^{Asn137}, is predominantly occupied by oligomannosidic glycans. Among AOC1 proteins, this N-glycosylation site has the highest degree of conservation

(publication II). Additionally, it is also the most conserved N-glycosylation site among the CAO proteins of the dataset in publication I. Altogether, the data suggests that this glycosylation site is conserved for about 450 million years (Venkatesh et al., 2014).

The mutation of Asn110 to a Gln in hAOC1, which prevents the N-glycosylation of the site, showed that the mutant activity was at least 100-fold lower than the wild type. The results showed that the oligomannose glycan attached to Asn110 is essential for hAOC1 folding and secretion. Based on the analysis of recombinant AOC1s, and hAOC3 from different sources this conserved N-glycosylation site consistently has oligomannosidic glycans (publication II). To analyse the relationship between the 3D structural features and the uncommon oligomannosidic glycosylation, the crystal structures of hAOC1 and hAOC3 as well as the homology model of pAOC1 were analysed in publication II. The overall sequence identity between hAOC1 and hAOC3 is 38%, while the sequence identity in the area surrounding the conserved N-glycosylation site its 67% (sequence similarity is 80%) (Figure 14A). In both proteins, the site is located in a cleft surrounded by hydrophobic residues in one side (Figure 14B-E). It also seems that this site is poorly accessible for the glycan processing enzymes that process high-mannose glycans to complex glycans, as the rate of final cis-Golgi mannose trimming to Man5 was decreased. Man5 is crucial for the addition of GlcNAc by the N-acetyl-glucosamiyltransferase-1 (GnT1), followed by the trimming to Man3 by alpha-mannosidae II. Thus the conversion to complex glycans seems to be limited. This is supported by the unsuccessful attempts to release the glycan at this site by N-glycosidase F, as the core region is not accessible (publication II).

Based on the structural analysis, the attached N-glycans are stabilized by conserved hydrogen bonds between the amide oxygen of the glycosylated Asn and the hydroxyl group of Thr or Ser in the N-glycosylation motif of hAOC1 or hAOC3, respectively. Furthermore, the conserved Gln residues form hydrogen bonds with the first GlcNAc of the glycan in each protein (Figure 14F-G). The N-acetyl methyl group of the second GlcNAc in the core trisaccharide forms hydrophobic interactions with two Phe residues that are part of the hydrophobic core of the D2 domain (Figure 15A). hAOC1 and

pAOC1 share a sequence identity of 85% and the three conserved glycosylation sites, including hAOC1^{Asn110}. In the model of porcine, this site is also located in a hydrophobic core with a similar accessibility than hAOC1^{Asn110} (publication II). Furthermore, the conserved hAOC1^{Asn110} N-glycosylation site seems to be essential for the secretion and for to the correct folding of both hAOC1 and likely also for CAOs deriving from different sources.

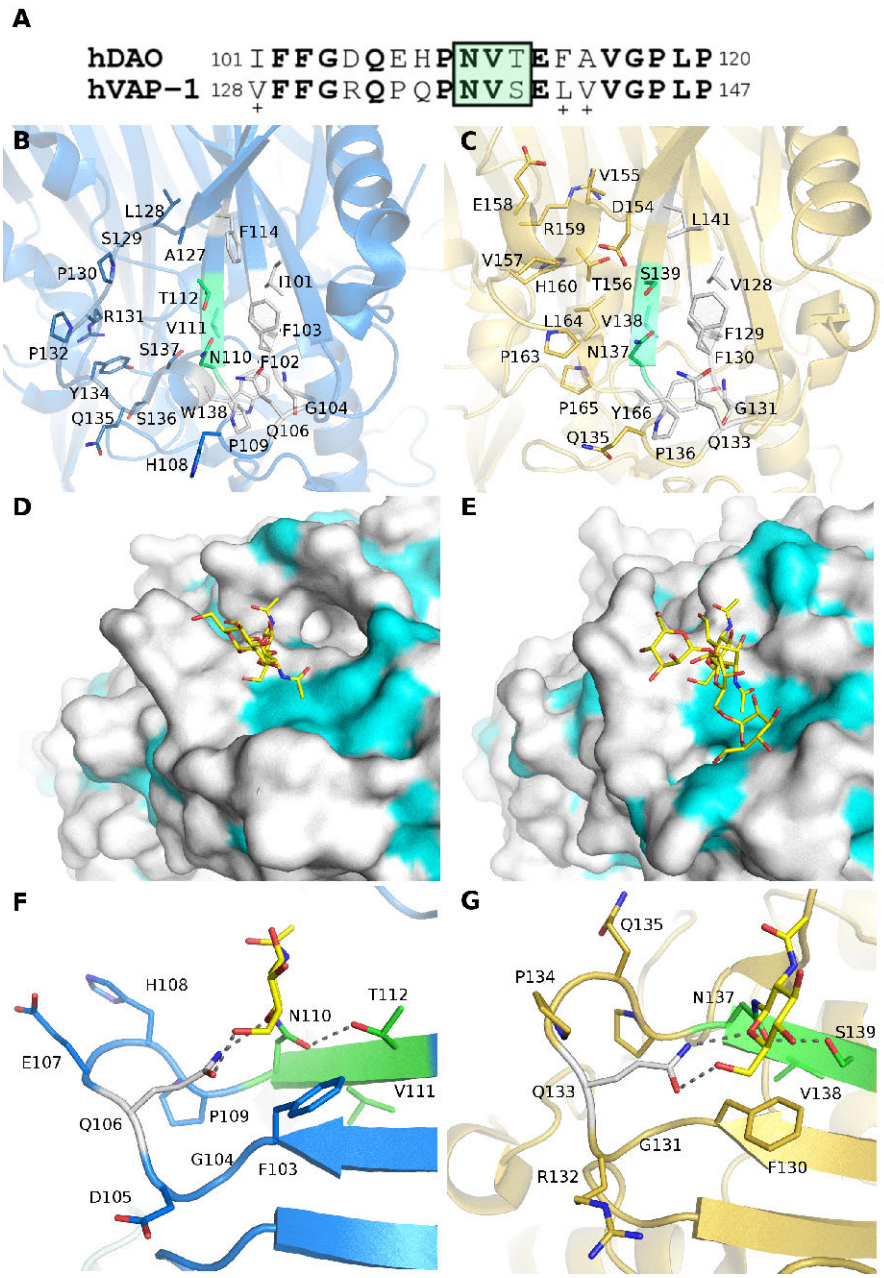


Figure 14. The conserved N-glycosylation site in hAOC1 and hAOC3. A: Sequence alignments with residues surrounding the Asn-Val-Ser/Thr motif. The identical residues are shown in bold, and the residues that preserve the hydrophobic nature marked with +. B: Close-up view of hAOC1^{Asn110}. C: Close-up view of hAOC3^{Asn137}. D: Surface view of hAOC1^{Asn110}. E: Surface view of hAOC3^{Asn137}. Conserved residues

and residues with a conserved hydrophobic nature shown as white sticks. The hydrophobic regions on the surface are in cyan. F: Interactions between the glycan and hAOC1 in the region surrounding $hAOC1^{Asn110}$. G: Interactions between the glycan and hAOC3 in the region surrounding $hAOC3^{Asn137}$ (Figure from publication II).

5.7. Structural characterization of the N-glycosylation sites in AOC1

The mutation of the remaining N-glycosylation sites, Asn538, Asn745 and Asn168, had an effect on hAOC1 secretion and reduced secretion by 71%, 32% and 13%, respectively (publication II). Thus, the structure of the surrounding regions in the remaining N-glycosylation sites of hAOC1 was also analysed in publication II to infer the possible structural role of these N-glycosylations. Asn168 is conserved only in primates and the composition of the glycans in this site varies between samples. This site has predominantly complex type glycans with high antennary and 100% fucosylation. Asn168 is located on a flat surface in the 3D structure of hAOC1 (Figure 15B). Consistently with the complex type of Asn168-attached glycans, Asn168 is readily accessible for the glycan-processing enzymes (publication II).

Asn538 and Asn745 are both occupied by similar complex-type glycans with similar antennary. They are both located in the D4 domain of hAOC1 and conserved in pAOC1. Asn538 is located next to the tip of Arm 1 of the β-hairpin (β4.3 β-strand), which extend from one monomer to the other. The region formed by the residues +3 to +7 from Asn538 is part of the tip of Arm 1 (Figure 15C). This evolutionary conserved region plays an important role in attaching Arm 1 into the D3 domain of the other monomer and the π-stacking interaction between Trp643 and Trp200 of the D3 domain seems to be particularly important. The inter-monomeric interactions are further stabilized by the hydrophobic interactions between the first N-acetyl methyl of the GlcNac with Pro606 and Trp609 from the D4 domain of the other monomer. Asn745 is located on the active site entrance near the C-terminus in the D4 domain. It is part of an extensive hydrogen-bonding network that connects the C-terminus to the core of D4 and to the Arm 2 from the other monomer (Figure

15D). The first GlcNAc of the glycan makes a water mediated hydrogen bond to Asn724 of the Arm 2 from the other monomer. Altogether, the location of the N-glycosylated sites on the 3D structure and the formed protein-glycan interactions explain the effect of the mutations on hAOC1 secretion and activity. In particular, the N-glycan attached to the site in the vicinity of the inter-domain Arms are important for the structural integrity, which is in accordance with the Arms being important for the maintenance of the dimer stability (Klema and Wilmot, 2012). pAOC1 has two glycosylation sites, which are not conserved with hAOC1, pAOC1^{Asn437} and pAOC1^{Asn749}. They form a cluster of glycans near the active site entrance together with pAOC1^{Asn794} (hAOC1^{Asn745}) (publication II). This cluster of glycans might facilitate local folding and dimerization efficiency.

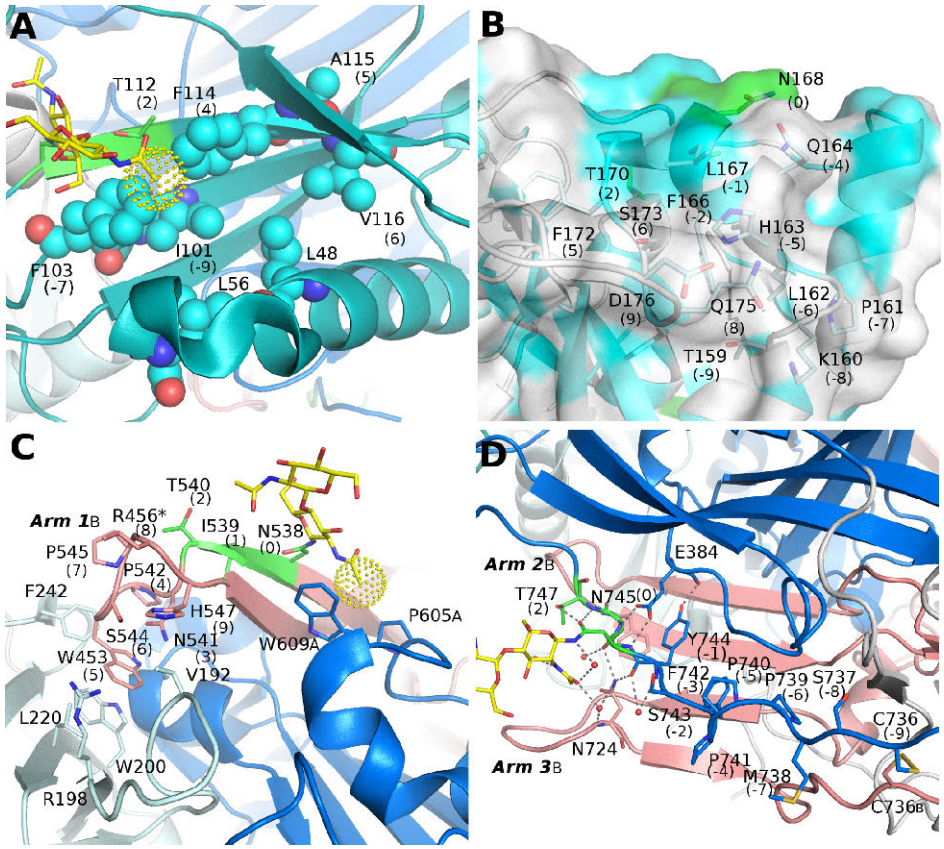


Figure 15. Close-up viewed of the N-glycosylation sites in AOC1 showing the surrounding residues and interactions. A: hAOC1^{Asn110}, B: hAOC1^{Asn168*}, C: hAOC1^{Asn538} D: hAOC1^{Asn745}. *not glycosylated in the published structures of hAOC1 (Figure from publication II).

5.8. Interactions between the Siglec-9 peptide and the N-glycosylated hAOC3

hAOC3 is a target for drug discovery and development and it also has a potential to be used as a biomarker for inflammation (Autio et al., 2013). The hAOC3-Siglec-9 interaction provides a powerful tool for *in vivo* imaging of leucocyte trafficking at sites of inflammation (Figure 16). In 2011, Aalto et al. found that Siglec-9 binds to the enzymatic grove of AOC3 and docking studies confirmed that the peptide indeed could fit into the active site cavity of AOC3 (Aalto et al., 2011). It was proposed that one of the arginines in the peptide

would bind covalently to TPQ but the later work with the extracellular part of Siglec-9 challenged the assumption (Elovaara et al., 2016). Although the *in vivo* imaging concept has been verified in various preclinical tests (Aalto et al., 2011; Jaakkola et al., 2000; Jensen et al., 2017; Li et al., 2013; Virtanen et al., 2015), details on the molecular level of the interaction still remain to be discovered. In publication III, the binding mode of the Siglec-9 peptide to hAOC3 was re-evaluated in light of the new data. Additionally, the N-glycans attached to hAOC3 (publication III) were for the first time taken into account as the N-glycans of hAOC3 were known to be involved in lymphocyte adhesion and in regulating the enzymatic activity (Maula et al., 2005).

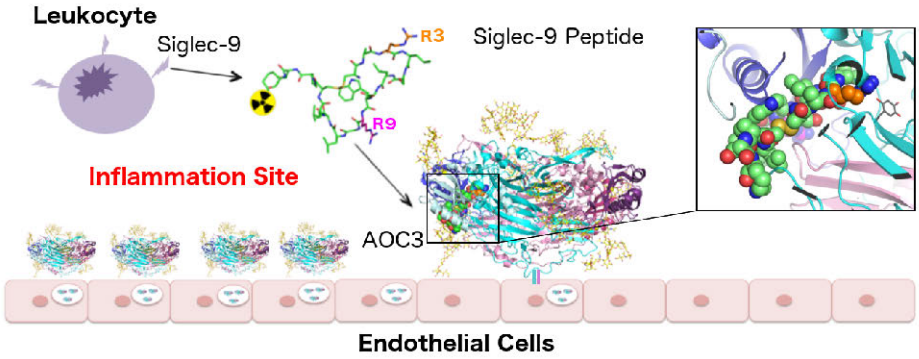


Figure 16. hAOC3 is stored inside the endothelial cells, where upon inflammatory stimuli it is released to the cell surface. A radiolabel peptide derived from the leucocyte ligand of hAOC3, Siglec-9, can be used for *in vivo* PET imaging of inflammation.

The docking results shows that the Siglec-9 peptide makes extensive intermolecular interactions with hAOC3 (Table 2) and effectively occupies its active site cavity (Figure 17A). R3 from the peptide protrudes in the active site forming hydrogen bonds with Tyr372 and Asn470, a salt bridge with Asp386 and hydrophobic interactions with Tyr384 (Figure 17B). R9 interacts in a pocked formed by hAOC3 and the glycan attached to Asn232. It forms hydrogen bonds to the first GlcNAc unit of Asn232 and interacts with Val209, Phe238 and Tyr488 from hAOC3 (Figure 17C). The other peptide-hAOC3 interactions of are listed in Table 3. The docking results are in accordance with Aalto *et al*. 2011, where R3 protrudes in the active

site. However, R3 does not directly bind to TPQ since based on the docking the binding of R3 in the catalytic site seems to occur when TPQ is in on-copper conformation. The fact that the peptide seems to bind to the on-copper TPQ conformation is in agreement with the earlier results that Siglec-9 is not a substrate for hAOC3 (Elovaara et al., 2016) as substrates bind to CAOs when TPQ is on the off-copper conformation (Klema and Wilmot, 2012) .

Next, the effect of the hAOC3 inhibitors, semicarbazide and imidazole, in the Siglec-9 peptide-hAOC3 interaction were studied (publication III). The predicted binding mode of the peptide was compared with docked semicarbazide and with the hAOC3 crystal structures in complex with imidazole (Figure 17D-F). Semicarbazide is an irreversible inhibitor of hAOC3, while imidazole reversibly inhibits hAOC3. Both inhibitors bind to TPQ but imidazole has an additional binding site in the active site channel (Figure 17D-E). The binding of the Siglec-9 peptide was not significantly affected when semicarbazide was injected before the peptide, suggesting that R3 can bind in a different binding site when the access to TPQ is blocked by semicarbazide. This site is likely the second binding site for imidazole (Figure 17E). This suggests that R3 interacts with TPQ like semicarbazide when TPQ is accessible but when the access to the active site is blocked by the bound semicarbazide, R3 may interact with Tyr394 and Thr212 like imidazole.

Similarly to the extracellular domain of Siglec-9 that increased the hAOC3 activity (Elovaara et al., 2016), the Siglec-9 peptide increased the activity of hAOC3 towards benzylamine by 1.3 fold (publication III). Since peptides with R3A, R9A and R3A/R9A mutations also increased the activity by approximately 1.4, 1.3 and 1.6 fold, respectively (publication III) the arginine residues in the peptide cannot be crucial for modulating hAOC3 activity. Results suggest that the Siglec-9 peptide would enhance the catalytic activity of hAOC3 by binding to TPQ when it is in the off-copper conformation but without blocking the access to TPQ (publication III). This could be the consequence of a conformational change in hAOC3, which allows a better access to TPQ. In a study made with lentil CAO it was showed that the both the D2 and D3 domains can move

(Dainese et al., 2014). Similar domain movement could occur in hAOC3. Furthermore, the results obtained with the Siglec-9 peptide are similar with the results obtained with the extracellular region of Siglec-9, suggesting a similar binding mode and further demonstrating that peptides are a good model to study protein-protein interactions.

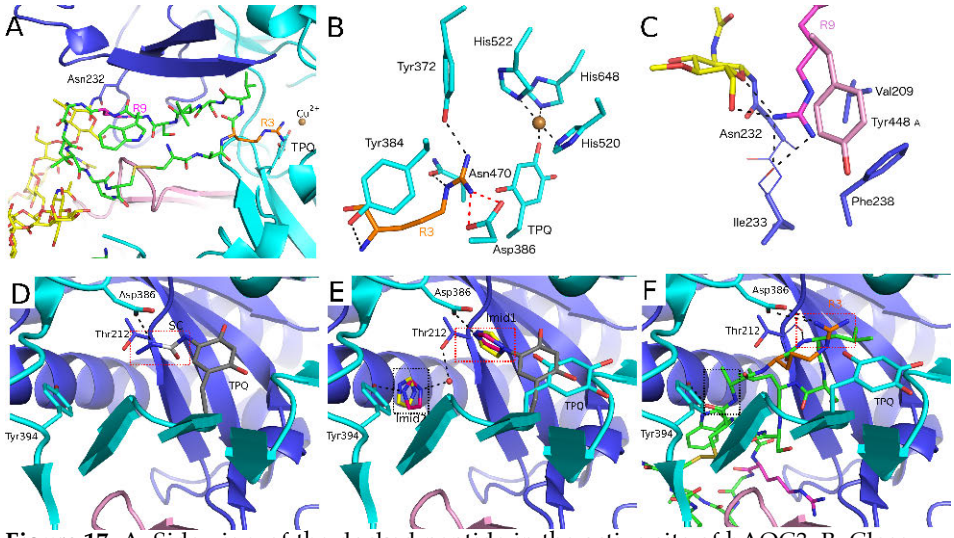


Figure 17. A: Side view of the docked peptide in the active site of hAOC3. B: Close-up view of R3 in the docked peptide. C: Close-up view of R9 in the docked peptide. D: Close-up view of docked semicarbazide (SC) with red dotted box showing the semicarbazide binding site. E: Close-up view of the imidazole binding sites Imi1 (red dotted box) and Imi2 (black dotted box). F: Peptide docked in the active site of hAOC3 with the sites Imi1 (red dotted box) and Imi2 (black dotted box) highlighted (Figure adapted form publication III).

Table 2. Predicted interactions between hAOC3 and the Siglec-9 derived peptide (sequence: CARLSLSWRGLTLCPSK). The residues in the active site of different organism used for the PET studies are also listed in the table, with conserved residues in bold. Residues in grey background are not present in the Siglec-9 $C2_2$ domain. SC = Side chain; MC = Main chain; H-bond = Hydrogen bond. [a]Arm I of chain A (Table from publication III).

Siglec-9 Peptide (Protein)	hAOC3 Chain B	ratAOC3	mouseAOC3	rabbitAOC3	pigAOC3	Predicted type of peptide- hAOC3 interaction
Residues						
C1	Tyr394	**Tyr**	**Tyr**	**Tyr**	**Tyr**	MC-SC H-bond
A2 (283)	Tyr176	**Tyr**	**Tyr**	**Tyr**	**Tyr**	Hydrophobic
R3 (284)	Tyr384 Asp386 Asn470 Tyr372	**Tyr** **Asp** **Asn** **Tyr**	**Tyr** **Asp** **Asn** **Tyr**	**Tyr** **Asp** **Asn** **Tyr**	**Tyr** **Asp** **Asn** **Tyr**	MC-SC H-bond Hydrophobic Salt bridge SC-SC H-bond SC-SC H-bond
L4 (285)	Tyr372 Thr212 Thr213	**Tyr** **Thr** **Thr**	**Tyr** **Thr** **Thr**	**Tyr** **Thr** Ala	**Tyr** **Thr** Ser	Hydrophobic MC-MC H-bond Hydrophobic
S5 (286)	Asn470	**Asn**	**Asn**	**Asn**	**Asn**	SC-MC H-bond
L6 (287)	Thr212 Phe227	**Thr** **Phe**	**Thr** **Phe**	**Thr** **Phe**	**Thr** **Phe**	Hydrophobic Hydrophobic
W8 (289)	Thr210 Leu177 Asp180	Lys Gln Gln	**Thr** Gln Glu	Val **Leu** **Asp**	**Thr** **Leu** **Asp**	MC-MC H-bond Hydrophobic Hydrophobic
R9 (290)	Val209 Ile233 Phe238 Tyr448[a]	Leu Leu **Phe** **Tyr**	Leu Leu **Phe** **Tyr**	Ala **Ile** **Phe** His	**Val** **Ile** **Phe** Arg	Hydrophobic SC-MC H-bond π-stacking π-stacking
P15 (296)	His762	Tyr	Tyr	**His**	**His**	MC-SC H-bond
S16 (297)	Ser419	**Ser**	**Ser**	**Ser**	**Ser**	MC-SC H-bond

5.9. Siglec-9 peptide binding in model organism

It is currently well known that the ligand binding and detailed catalytic properties of CAOs differ between species (Dunkel et al., 2011), e.g. mouse AOC3 can oxidize a wider range of substrates than hAOC3 (Bono et al., 1999). Therefore, results obtained with one species cannot be directly applied to other species without a careful comparison of the binding site. The preclinical PET studies with the Sigelc-9 peptide targeting AOC3 have been made using many model organisms (Aalto et al., 2011; Jaakkola et al., 2000; Jensen et al., 2017; Li et al., 2013; Virtanen et al., 2015). Thus, in publication III the species-specific differences of AOC3 that could influence the binding of the Siglec-9 peptide to AOC3 in the model organisms were taken into account.

Siglec-9 peptide is a bulky ligand that interacts in the least conserved part of the active site of the AOC3 protein. The overall sequence identities of hAOC3 are 82% and 83% to mouse and rat AOC3s. The pig and rabbit enzymes share a slightly higher identity

to hAOC3 (85% and 84%, respectively). The comparison of the residues predicted to be involved in peptide binding show that in general these residues are conserved or conservatively replaced (Table 2). R3, which binds near TPQ, interacts with the most conserved part of the active site channel. All the residues interacting with R3 are conserved and, thus, we could expect similar interaction in all the above-mentioned species. R9 interacts with the least conserved part of the channel. In fact, R9 interacts with two non-conserved residues, Val209 and Tyr448. Val209 is conserved in pAOC3, but replaced by Ala in rabbit and Leu in mouse and rat enzymes. Tyr448 is replaced by a His in rabbit and Arg in pig AOC3. Moreover, rabbit AOC3 lacks the N-glycosylation site corresponding to hAOC3^{Asn232} and thus it cannot make the glycan-R9 interactions. In accordance with the residue differences, the Siglec-9 peptide has been reported to bind pig and rabbit AOC3 more weakly compared to human AOC3 (Jensen et al., 2017). Altogether, these results underline the importance of the acknowledging the species-specific differences of AOC3 in model organisms and its consequences for *in vivo* imaging of inflammation using the Siglec-9 peptide.

5.10. Characterization of SynPAO

In publication IV, a cyanobacteria PAO (*Synechocystis sp*. PCC 6803 PAO; SynPAO) was characterized for the first time. SynPAO has substrate preference for spermine and spermidine. It can weakly oxidize N^1-acetyl spermine but it has no activity towards the diamines putrecine and cadaverine. The analysis of the reaction products (publication IV) showed that SynPAO is involved in the back-conversion pathway (*exo*-mode oxidation) similar to AtPAO1-5 (Fincato et al., 2011; Tavladoraki et al., 2016, 2006).

SynPAO shares sequence identities from ~18% to 50% with the PAO sequences studied in publication IV and ~50% sequence identity with the cyanobacterial PAOs. The phylogenetic tree shows three major branches of PAOs (Figure 18): branch I with animal PAOs, AtPAO5 and similar proteins, branch II with cyanobacterial, bacterial and plant PAOs (including SynPAO, AtPAO1 and ZmPAO), and branch III with AtPAO2, AtPAO3, AtPAO4 and

similar proteins. The further analysis was focused on branch II, which has SynPAO. This branch is divided into bacterial and plant sub-branches, supported with a bootstrap value of 100% (Figure 18). The bacterial branch includes SynPAO and other bacterial proteins that are not yet characterized. The plant branch is divided into two sub-branches. One of them has proteins involved in the terminal catabolic pathway (the TC branch) and the other branch has proteins involved in the back-conversion pathway (the BC branch) (Figure 18). The TC branch contains ZmPAO, HvPAO1, HvPAO2 and OsPAO7 known to be involved in the terminal catabolic pathway (Cervelli et al., 2006; Liu et al., 2014b; Tavladoraki et al., 1998), Additionally, the TC branch has the PAO2 proteins that form their own sub-branch (bootstrap 100%). These proteins have a C-terminal extension with a highly conserved AEAK sequence motif, which is a vacuolar sorting signal necessary for the intracellular location (Cervelli et al., 2006). Furthermore, the branching pattern is in accordance with the substrate preference of the barley proteins HvPAO1 and HvPAO2 (Cervelli et al., 2006). AtPAO1 that is involved in the back-conversion pathway and the *Nicotina tabacum* PAO (NtPAO) that has 83% sequence identity to AtPAO1 are in the BC branch. Accordingly, it has been suggested that NtPAO and AtPAO1 have similar catalytic properties (Tavladoraki et al., 2006). A *Malus domestica* (MdPAO) also shares the BC branch and it is likely involved in the back conversion pathway. The phylogenetic tree in Figure 18 has a topology similar with the previously published trees of plant PAOs (Ono et al., 2012; Takahashi et al., 2010; Tavladoraki et al., 2016) and the tree topology is consistent with the known oxidation modes of plant PAOs. Furthermore, the topology is in accordance with the findings made by Ahou and co-workers (Ahou et al., 2014), in which AtPAO5 was found to be more similar to mouse spermine oxidase (MmSMO) and mouse polyamine oxidase (MmPAO) than to AtPAO1-4.

Spermine was docked into the 3D model of SynPAO to predict the binding mode and the results were compared with the spermine binding mode of the PAOs having different oxidation modes (Figure 19). Additionally, sequences from the proteins in branch II (Figure 18) were compared with the active site of the spermine complex

structures of ZmPAO (Fiorillo et al., 2011) (terminal conversion) and yeast Fms1 (Huang et al., 2005) (back conversion), the model of MmSMO (Tavladoraki et al., 2011), and the 3D model of SynPAO (publication IV) with docked spermine (Table 3). The aromatic sandwich formed by Phe403 and Tyr439 in ZmPAO (Figure 19A) is important for creating the correct substrate conformation for the *endo*-mode reaction (Binda et al., 1999; Fiorillo et al., 2011). This aromatic sandwich is also conserved in plant PAOs involved in the *exo*-mode reaction (Table 3, blue) and, thus, it is not the key for the *endo*-mode oxidation. Similar to Fms1 and MmSMO, bacterial PAOs lack this aromatic sandwich (Figure 19). Tyr439 in ZmPAO is replaced by a non-aromatic residue (Table 3, Thr440 in SynPAO) and Phe403 of ZmPAO is replaced by a Tyr in bacterial PAOs, MmSMO, and Fms1 (Table 4, Tyr403 in SynPAO). In ZmPAO, Glu62 and Glu170 are important in positioning spermine (Binda et al., 2001, 1999). These residues are totally conserved in all plant PAOs that catalyse the *endo*-mode reaction (Table 3). Glu62 (ZmPAO) is either a Ala or Val in plant PAOs catalysing the *exo*-mode of reaction (Table 3). These PAOs however have a conserved Glu that could replace this residue (N195 Fms1 Table 3). In Fms1, MmSMO and bacterial PAOs, Glu62 of ZmPAO corresponds to a His and the corresponding residue is Gln94 in SynPAO (Table 3). Glu170 of ZmPAO is conserved in all the studied plant and bacterial PAOs (Table 3). Mutational studies of Fms1 show that His76 positions the substrate for the hydride transfer to flavin (Adachi et al., 2012). Accordingly, the mutation H82Q in MmSMO increased the K_m about 2-fold compared with wild-type enzyme supporting the role of this residue in substrate binding (Tavladoraki et al., 2011). Based on docking studies in publication IV the corresponding residue in SynPAO, Glu62, is predicted to interact with the N^1 nitrogen of spermine (Figure 19C). Moreover, the docking results are consistent with the *exo*-oxidation mode oxidized by SynPAO. The phylogenetic study showing that SynPAO shares the same sub-tree as the other proteins known to be involved in the back-conversion pathway are in line with the experimental results. Moreover, Gln94, Tyr403, and Thr440 were identified as the key residues in the active site of SynPAO (Figure 19C).

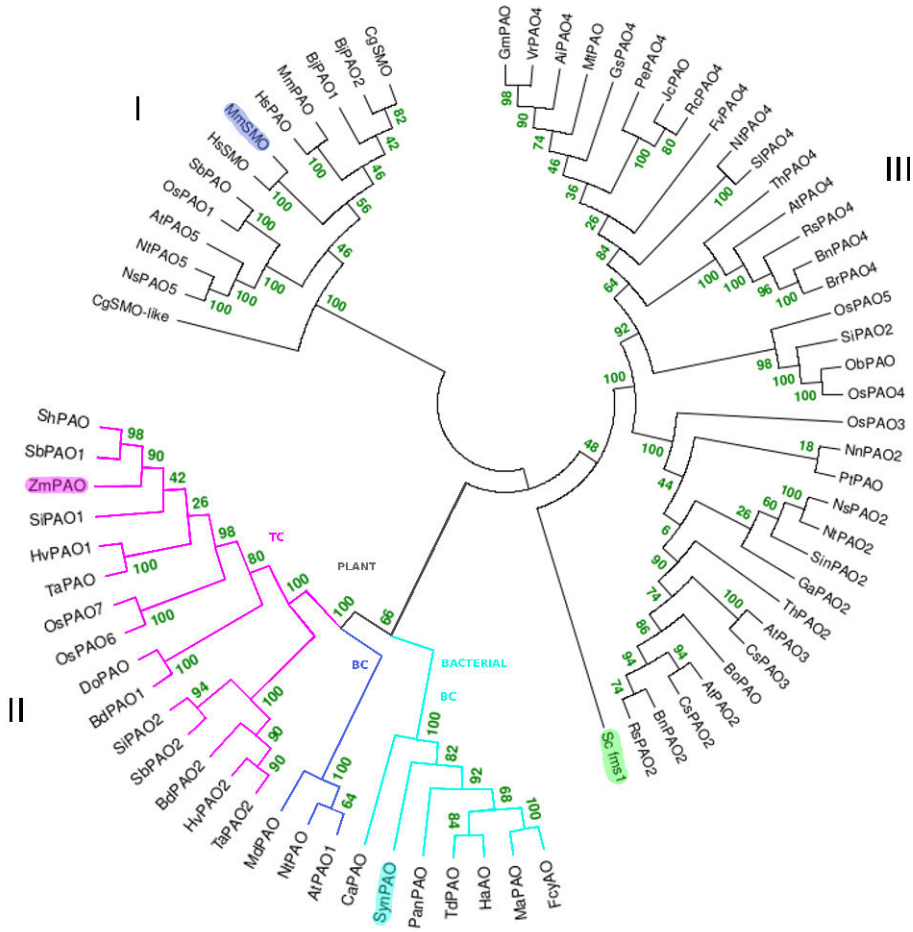


Figure 18. ML tree of amine oxidases from different organisms. Branch II contains SynPAO (highlighted in cyan) together with other cyanobacteria and bacterial proteins within the bacterial subbranch (cyan). In the plant subbranch there are some proteins known to catalyze the terminal catabolism (TC), including ZmPAO (highlighted in pink), HvPAO1, HvPAO2 and OsPAO7, in the pink subbranch. The blue branch of the plant subbranch has plant proteins that catalyze the back-conversion pathway (BC) e.g. AtPAO1. Yeast Fms1 is highlighted in green. To identify the species origin of the proteins, the following species acronyms were used: Ai, *Arachis ipaensis*; At, *Arabidopsis thaliana*; Bd, *Brachypodium distachyon*; Bj, *Branchiostoma japonicum*; Bn, *Brassica napus;* Bo, *Brassica oleracea*; Br, *Brassica rapa*; Ca, *Cyanobacterium aponinum*; Cg, *Crassostrea gigas*; Cs, *Camelina sativa*; Do, *Dichanthelium oligosanthes*; Fcy, *Filamentous cyanobacterium*; Fv, *Fragaria vesca*; Ga, *Gossypium arboreum*; Gm, *Glycine max*; Gs, *Glycine soja*; Ha, *Halomonas alkaliantartica*; Hs; *Homo sapiens*; Hv, *Hordeum vulgare*; Jc, *Jatropha curcas*; Ma, *Microcystis aeruginosa*;

Md, *Malus domestica*; Mm, *Mus musculus*; Mt, *Medicago truncatula*; Nn, *Nelumbo nucifera*; Ns *Nicotiana sylvestris*; Nt, *Nicotiana tabacum*; Ob, *Oryza brachyantha*; Os, *Oryza sativa*; Pan, *Pseudanabaena sp*. PCC 6802; Pe, *Populus euphratica*; Pt, *Populus trichocarpa*; Rc, *Ricinus communis*; Rs, *Raphanus sativus*; Sb, *Sorghum bicolor*; Sc, *Saccharomyces cerevisiae*; Sh, *Saccharum hybrid*; Si, *Setaria italica*; Sin, *Sesamum indicum*; Sl, *Solanum lycopersicum*; Syn, *Synech*ocystis; Ta, *Triticum aestivum*, Td, *Thiothrix disciformis*; Th, *Tarenaya hassleriana*; Vr, *Vigna radiata*; Zm, *Zea mays*. The bootstrap values supporting each branch are in green color and a bootstrap value above 70% provides good support for the branch. (Figure from publication IV).

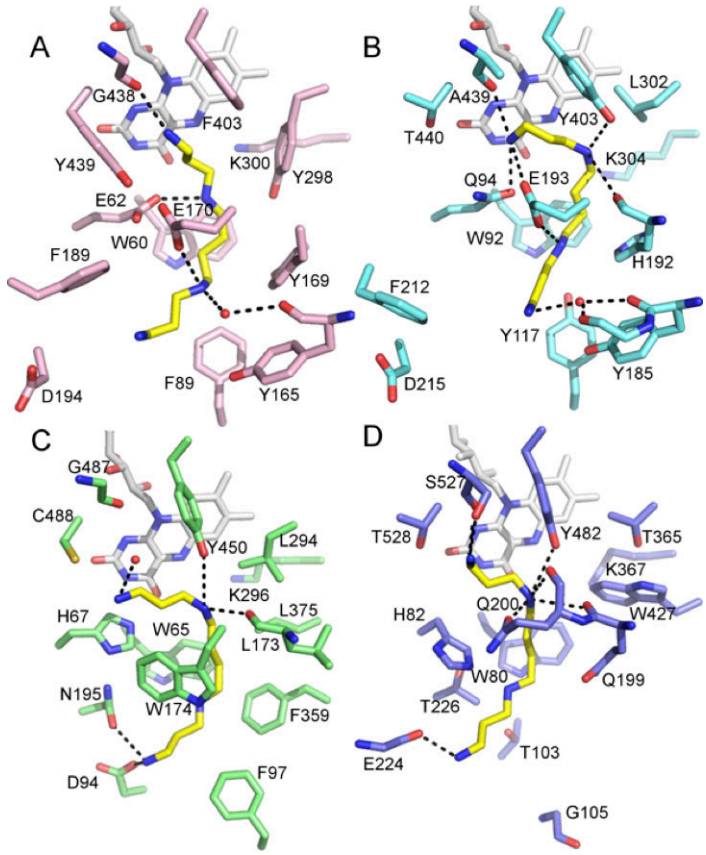


Figure 19. Comparison of the spermine binding mode of PAOs with different oxidation mode. (A) The X-ray structure of ZmPAO in complex spermine (PBD ID: 3KU9, M300K mutated). (B) The homology model of SynPAO with the docked spermine. (C) The X-ray structure of yeast Fms1 in complex with spermine (PDB ID: 1XPQ; chain A). (D) The homology model of MmSMO with the docked spermine. FAD is shown in white sticks and spermine in yellow sticks (Figure 6 in publication IV).

Table 3. Residues from the Spermine binding site of PAOs that share branch II (Figure 18) with SynPAO compared with the residues in Fms1 and MmSMO. The highly conserved positions in plant and bacterial PAOs have orange and grey backgrounds, respectively. Residues positioning spermine are highlighted in yellow. The totally conserved residues in the branch II proteins (Figure 16) are shown in bold (Table from publication IV).

Protein	Residues													
ZmPAO	**W60**	E62	F89	Y165	Y169	**E170**	F189	D194	V196	Y298	**K300**	F403	G438	Y439
SbPAO1/ShPAO/DoPAO	**W**	E	F	Y	Y	**E**	F	D	V	Y	**K**	F	G	Y
TaPAO/HvPAO1	**W**	E	F	F	Y	**E**	T	D	V	Y	**K**	F	G	Y
SiPAO1	**W**	E	F	Y	Y	**E**	F	D	T	Y	**K**	F	G	Y
OsPAO7	**W**	E	F	F	Y	**E**	F	D	N	Y	**K**	F	G	Y
OsPAO6	**W**	E	F	F	Y	**E**	F	D	T	Y	**K**	F	G	Y
BdPAO1	**W**	E	F	Y	F	**E**	F	D	V	Y	**K**	F	G	Y
TaPAO2	**W**	E	F	F	Y	**E**	F	D	A	Y	**K**	Y	G	Y
BdPAO2	**W**	E	F	F	Y	**E**	A	E	N	Y	**K**	Y	G	Y
SbPAO2	**W**	E	F	F	Y	**E**	N	E	N	Y	**K**	Y	G	Y
SiPAO2	**W**	E	F	Y	Y	**E**	N	E	S	Y	**K**	Y	G	Y
HvPAO2	**W**	E	F	Y	Y	**E**	T	D	A	Y	**K**	Y	G	Y
AtPAO1/NtPAO	**W**	A	Y	I	F	**E**	Y	E	E	Y	**K**	Y	G	Y
MdPAO	**W**	V	Y	T	F	**E**	F	E	E	Y	**K**	Y	G	Y
SynPAO	**W92**	Q94	Y117	Y185	H192	**E193**	F212	D215	A217	L302	**K304**	Y403	A439	T440
CaPAO	**W**	H	S	L	Q	**E**	F	D	S	L	**K**	Y	A	T
PanPAO	**W**	H	Y	I	H	**E**	F	E	A	L	**K**	Y	G	S
FcyAO	**W**	H	Y	I	H	**E**	F	D	V	L	**K**	Y	G	T
MaPAO	**W**	H	Y	I	H	**E**	F	N	D	L	**K**	Y	G	T
TdPAO	**W**	H	Y	V	Q	**E**	F	D	V	L	**K**	Y	G	T
HaAO	**W**	H	Y	C	Q	**E**	F	D	V	L	**K**	Y	G	T
Fms1	**W65**	H67	F97	Y170	L173	W174	H191	G193	N195	L294	**K296**	Y450	G487	C488
MmSMO	**W80**	H82	G105	L195	Q199	Q200	L219	E224	T226	T364	**K376**	Y482	S527	T528

Residues in ZmPAO and Fms1 are numbered according to the crystal structures. Residues in SynPAO are numbered according to the sequence entry (UniProt Q6ZEN7) and MmSMO according to (Tavladoraki et al., 2011).

6. Conclusion

This thesis describes the use of bioinformatics methods to characterize proteins from the amine oxidase family. Sequence analysis, molecular docking, homology modelling and phylogenetics were used as the methods and complemented with the results from the experimental studies to increase the reliability of the computational results.

In publication I, evolutionary studies, sequence and structural data were used together to show that mammalian CAOs can be classified based on the active site motif (X1-X2-Asn-Tyr-Asp). X2 can be used to discriminate between AOC1 (X2=Tyr), AOC2 (X2=Gly) and AOC3/4 (X2=Leu), whereas residue X1 can be used to discriminate between AOC3 (X1=Leu) and AOC4 (X1=Met). The classification is consistent with the substrate preferences of these proteins. The revised classification rules can be used to correctly classify new CAO sequences and aid in the reclassification of wrongly annotated sequences. Additionally, a third residue that seems to play a role in the substrate preference was found in the active site of mammalian CAOs, Asp186 in hAOC1, His206 in hAOC2, Thr212 in hAOC3, and Asn211 in bAOC4. Moreover, publication I also shows the species-specific expression profile of CAOs, which means that not all mammals have the same number of CAOs and that when studying one specific CAO it is important to be aware of the other CAOs expressed in that organism.

Publication II studies the N-glycosylation sites of hAOC1, in particular the N-glycosylation site conserved in mammalian CAOs ($hAOC1^{Asn110}$). This conserved site is occupied by oligomannosidic glycans in hAOC1, pAOC1 and hAOC3. In these proteins, this site is located in a hydrophobic cleft that interacts with the core trisaccharide of the attached N-glycan. This site is the key for hAOC1 segregation, as it seems to be important for the correct folding of the D2 domain and, consequently, the folding of a functional hAOC1. Based on the conservation, it might have a similar role in mammalian CAOs. The other N-glycosylation sites in

hAOC1, located at Asn538 and Asn745, also seem to have an important structural role, stabilizing the inter-monomeric arms of the D4 domain. Understanding the role of the carbohydrates for the half-life and clearance of hAOC1 is important, if recombinant AOC1 is to be used in patients to degrade exogenous and endogenous histamine. For example, pAOC1 has been used in the treatment of several diseases with limited success (Laymon and Cumming, 1939). Knowing the glycan structures involved in hAOC1 clearance together with modern biotechnological methods, might help to circumvent the issues that arose decades ago when this therapy was first reported.

In publication III, the interaction mode between hAOC3 and the Siglec-9 derived peptide was re-evaluated in the light of the new data. Understanding of atomic level interactions as well as the species-specific differences in the residues involved in the interactions is important for optimizing the peptide for the PET studies. The results obtained with the Siglec-9 derived peptides are consistent with the previous studies using the full-length extracellular region of Siglec-9. Thus, this study demonstrates that analysis of peptide-protein interactions can function as a good model for studying protein-protein interactions. Moreover, the *Synechocystis sp*. PCC 6803 PAO (SynPAO) was characterized in publication IV. The phylogenetic analysis together with structural data (docking and structural comparison) identified the key residues that play a role in spermine oxidase in SynPAO.

The 3D coordinates alone give information about the shape of the protein, however, when relating them with other data, they provide valuable information on the type of fold and the functional classification of the protein. This demonstrates the importance of using an integrative approach of data analysis, which by connecting different bits of information from different sources, allows us to obtain biologically meaningful data and expand the understanding about various aspects of proteins. Proteins from amine oxidase family are widely distributed in the tree of life, and in general there is more than one member of this family present in each organism. Here the members of the amine oxidase protein family were

characterized using different bioinformatics methods and the results were analysed together with experimental data. The new classification rules for mammalian CAOs allow classifying correctly CAOs into their subfamily, and aid in the design of new subfamily specific inhibitors. The work with CAOs also will benefit the design of *in vitro* tests to detect individual CAO levels, that could be used for diagnostic purposes and to improve the current CAO-targeted drugs and biomarkers. The other amine oxidase member studied here PAO, also has a variety of potential applications. The knowledge obtained here can be used to understand the residues behind the substrate preference in PAOs and for example, to guide protein engineering to develop biosensors to detect biogenetic amines. As proteins from this family are used both as drug targets and biomarkers, and their role in different diseases is widely studied, the research presented in this thesis can be used to further develop the research in the amine oxidase family. Moreover, this thesis also shows that when used correctly bioinformatics can be successfully used to optimize different research projects.

7.References

Aalto, K., Autio, A., Kiss, E.A., Elima, K., Nymalm, Y., Veres, T.Z., Marttila-Ichihara, F., Elovaara, H., Saanijoki, T., Crocker, P.R., Maksimow, M., Bligt, E., Salminen, T.A., Salmi, M., Roivainen, A., Jalkanen, S., 2011. Siglec-9 is a novel leukocyte ligand for vascular adhesion protein-1 and can be used in PET imaging of inflammation and cancer. Blood 118, 3725–3733.

Aalto, K., Maksimow, M., Juonala, M., Viikari, J., Jula, A., Kähönen, M., Jalkanen, S., Raitakari, O.T., Salmi, M., 2012. Soluble vascular adhesion protein-1 correlates with cardiovascular risk factors and early atherosclerotic manifestations. Arterioscler. Thromb. Vasc. Biol. 32, 523–532.

Abella, a., García-Vicente, S., Viguerie, N., Ros-Baró, a., Camps, M., Palacín, M., Zorzano, a., Marti, L., 2004. Adipocytes release a soluble form of VAP-1/SSAO by a metalloprotease- dependent process and in a regulated manner. Diabetologia 47, 429–438.

Adachi, M.S., Taylor, A.B., Hart, P.J., Fitzpatrick, P.F., 2012. Mechanistic and structural analyses of the roles of active site residues in yeast polyamine oxidase Fms1: Characterization of the N195A and D94N enzymes. Biochemistry 51, 8690–8697.

Ahou, A., Martignago, D., Alabdallah, O., Tavazza, R., Stano, P., Macone, A., Pivato, M., Masi, A., Rambla, J.L., Vera-Sirera, F., Angelini, R., Federico, R., Tavladoraki, P., 2014. A plant spermine oxidase/dehydrogenase regulated by the proteasome and polyamines. J. Exp. Bot. 65, 1585–1603.

Airenne, T.T., Nymalm, Y., Kidron, H., Smith, D.J., Pihlavisto, M., Salmi, M., Jalkanen, S., Johnson, M.S., Salminen, T. a., 2005. Crystal structure of the human vascular adhesion protein 1: Unique structural features with functional implications. Protein Sci. 14, 1964–1974.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Arndtz, K., Corrigan, M., Rowe, A., Kirkham, A., Barton, D., Fox, R.P., Llewellyn, L., Athwal, A., Wilkhu, M., Chen, Y.Y., Weston, C., Desai, A., Adams, D.H., Hirschfield, G.M., 2017. Investigating the safety and activity of the use of BTT1023 (Timolumab), in the treatment of patients with primary sclerosing cholangitis (BUTEO): A single-arm, two-stage, open-label, multi-centre, phase II clinical trial protocol. BMJ Open 7.

Autio, A., Jalkanen, S., Roivainen, A., 2013. Nuclear imaging of inflammation: homing-associated molecules as targets. EJNMMI Res. 3, 1. doi:10.1186/2191-219X-3-1

Benkert, P., Kunzli, M., Schwede, T., 2009. QMEAN server for protein model quality estimation. Nucleic Acids Res. 37, W510–W514.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The protein data bank. Nucleic Acids Res. 28, 235–242.

Bieberich, E., 2014. Synthesis, processing, and function of N-glycans in N-glycoproteins. Adv Neurobiol. 9, 47–70.

Binda, C., Angelini, R., Federico, R., Ascenzi, P., Mattevi, A., 2001. Structural bases for inhibitor binding and catalysis in polyamine oxidase. Biochemistry 40, 2766–2776.

Binda, C., Coda, A., Angelini, R., Federico, R., Ascenzi, P., Mattevi, A., 1999. A 30 Å long U-shaped catalytic tunnel in the crystal structure of polyamine oxidase. Structure 7, 265–276.

Bligt-Lindén, E., Arunchalam, R., Parkash, V., Salminen, T.A., 2012. Structural comparison of the active site channels in rodent and primate vascular adhesion protein-1. J Neural Transm 120.

Bligt-Lindén, E., Pihlavisto, M., Szatmári, I., Otwinowski, Z., Smith, D.J., Lázár, L., Fülöp, F., Salminen, T. a., 2013. Novel pyridazinone inhibitors for vascular adhesion protein-1 (VAP-1): Old target-new inhibition mode. J. Med. Chem. 56, 9837–9848.

Boehm, T., Pils, S., Gludovacz, E., Szoelloesi, H., Petroczi, K., Majdic, O., Quaroni, A., Borth, N., Valent, P., Jilma, B., 2016. Quantification of human diamine oxidase. Clin. Biochem.8–15.

Boffi, A., Favero, G., Federico, R., Macone, A., Antiochia, R., Tortolini, C., Sanzó, G., Mazzei, F., 2015. Amine oxidase-based biosensors for spermine and spermidine determination. Anal. Bioanal. Chem. 407, 1131–1137.

Bonaiuto, E., Lunelli, M., Scarpa, M., Vettor, R., Milan, G., Di Paolo, M.L., 2010. A structure-activity study to identify novel and efficient substrates of the human semicarbazide-sensitive amine oxidase/VAP-1 enzyme. Biochimie 92,858–68.

Bono, P., Jalkanen, S., Salmi, M., 1999. Mouse vascular adhesion protein 1 is a sialoglycoprotein with enzymatic activity and is induced in diabetic insulitis. Am. J. Pathol. 155, 1613–1624.

Boomsma, F., Bhaggoe, U.M., Van Der Houwen, A.M.B., Van Den Meiracker, A.H., 2003. Plasma semicarbazide-sensitive amine oxidase in human (patho)physiology. Biochim. Biophys. Acta - Proteins Proteomics 1647, 48–54.

Boomsma, F., Hut, H., Bhaggoe, U., Houwen, A. Van Der, Meiracker, A. Van Den, 2005. From Cell To Circulation. Heart Fail. 11, 122–126.

Bouchereau, A., Aziz, A., Larher, F., Martin-Tanguy, J., 1999. Polyamines and environmetal challenges: recent development. Plant Sci. 140, 103–125.

Cervelli, M., Bianchi, M., Cona, A., Crosatti, C., Stanca, M., Angelini, R., Federico, R., Mariottini, P., 2006. Barley polyamine oxidase isoforms 1 and 2, a peculiar case of gene duplication. FEBS J. 273, 3990–4002.

Cervelli, M., Polticelli, F., Angelucci, E., Di Muzio, E., Stano, P., Mariottini, P., 2015. Pacific oyster polyamine oxidase: A protein missing link in invertebrate evolution. Amino Acids 47, 949–961.

Choi, Y.H., Matsuzaki, R., Suzuki, S., Tanizawa, K., 1996. Role of conserved Asn-Tyr-Asp-Tyr sequence in bacterial copper/2,4,5- trihydroxyphenylalanyl quinone-containing histamine oxidase. J. Biol. Chem. 271, 22598–22603.

Choong, Y.S., Tye, G.J., Lim, T.S., 2013. Minireview: Applied structural bioinformatics in proteomics. Protein J. 32, 505–511.

Cona, A., Rea, G., Angelini, R., Federico, R., Tavladoraki, P., 2006. Functions of amine oxidases in plant development and defence. Trends Plant Sci. 11, 80–88.

Dainese, E., Sabatucci, A., Pintus, F., Medda, R., Angelucci, C.B., Floris, G., Maccarrone, M., 2014. Domain mobility as probed by small-angle X-ray scattering may account for substrate access to the active site of two copper-dependent amine oxidases. Acta Crystallogr. D. Biol. Crystallogr. 70, 2101–10.

Dawkes, H., 2001. Copper amine oxidase: cunning cofactor and controversial copper. Curr. Opin. Struct. Biol. 11, 666–673.

DeLano, W.L., 2002. The PyMOL Molecular Graphics System, Version 1.1.

Schrödinger LLC

Di Paolo, M.L., Stevanato, R., Corazza, A., Vianello, F., Lunelli, L., Scarpa, M., Rigo, A., 2003. Electrostatic compared with hydrophobic interactions between bovine serum amine oxidase and its substrates. Biochem. J. 371, 549–556.

Doherty, G.J., McMahon, H.T., 2008. Mediation, Modulation, and Consequences of Membrane-Cytoskeleton Interactions. Annu. Rev. Biophys. 37, 65–95.

Dunkel, P., Balogh, B., Meleddu, R., Maccioni, E., Gyires, K., Mátyus, P., 2011. Semicarbazide-sensitive amine oxidase/vascular adhesion protein-1: a patent survey. Expert Opin. Ther. Pat. 21, 1453–1471.

Eklund, E.A., Freeze, H.H., 2005. Essentials of glycosylation. Semin. Pediatr. Neurol. 12, 134–143. 1

Elmore, B.O., Bollinger, J. a, Dooley, D.M., 2002. Human kidney diamine oxidase: heterologous expression, purification, and characterization. J. Biol. Inorg. Chem. 7, 565–579.

Elovaara, H., Kidron, H., Parkash, V., Nymalm, Y., Bligt, E., Ollikka, P., Smith, D.J., Pihlavisto, M., Salmi, M., Jalkanen, S., Salminen, T. A., 2011. Identification of two imidazole binding sites and key residues for substrate specificity in human primary amine oxidase AOC3. Biochemistry 50, 5507–20.

Elovaara, H., Parkash, V., Fair-Mäkelä, R., Salo-Ahen, O.M.H., Guédez, G., Bligt-Lindén, E., Grönholm, J., Jalkanen, S., Salminen, T.A., 2016. Multivalent interactions of human primary amine oxidase with the v and C22 domains of sialic acid-binding immunoglobulin-like lectin-9 regulate its binding and amine oxidase activity. PLoS One 11, 1–22.

Ernberg, K., McGrath, A.P., Peat, T.S., Adams, T.E., Xiao, X., Pham, T., Newman, J., McDonald, I.A., Collyer, C.A., Guss, J.M., 2010. A new crystal form of human vascular adhesion protein 1. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66, 1572–1578.

Felsenstein, J., 1985. Confidence-Limits on phylogenies - an approach using the bootstrap. Evolution (N. Y). 39, 783–791.

Fincato, P., Moschou, P.N., Spedaletti, V., Tavazza, R., Angelini, R., Federico, R., Roubelakis-Angelakis, K.A., Tavladoraki, P., 2011. Functional diversity inside the Arabidopsis polyamine oxidase gene family. J. Exp. Bot. 62, 1155–1168.

Finney, J., Moon, H.J., Ronnebaum, T., Lantz, M., Mure, M., 2014. Human copper-dependent amine oxidases. Arch. Biochem. Biophys. 546, 19–32.

Fiorillo, A., Federico, R., Polticelli, F., Boffi, A., Mazzei, F., Di Fusco, M., Ilari, A., Tavladoraki, P., 2011. The structure of maize polyamine oxidase K300M mutant in complex with the natural substrates provides a snapshot of the catalytic mechanism of polyamine oxidation. FEBS J. 278, 809–821.

Foot, J.S., Deodhar, M., Turner, C.I., Yin, P., van Dam, E.M., Silva, D.G., Olivieri, A., Holt, A., McDonald, I. a, 2012. The discovery and development of selective 3-fluoro-4-aryloxyallylamine inhibitors of the amine oxidase activity of semicarbazide-sensitive amine oxidase/vascular adhesion protein-1 (SSAO/VAP-1). Bioorg. Med. Chem. Lett. 22, 3935–40.

Frébort, I., Peč, P., Luhová, L., Toyama, H., Matsushita, K., Hirota, S., Kitagawa, T., Ueno, T., Asano, Y., Kato, Y., Adachi, O., 1996. Two amine oxidases from Aspergillus niger AKU 3302 contain topa quinone as the cofactor: Unusual cofactor link to the glutamyl residue occurs only at one of the enzymes. Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol. 1295, 59–72.

Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., 2004. Glide: a new approach for rapid, accurate docking and scoring. 1. Method

and assessment of docking accuracy. J. Med. Chem. 47, 1750–1759.

Gaule, T.G., Smith, M. a., Pearson, A.R., Knowles, P.F., McPherson, M.J., 2015. Probing the Molecular Mechanisms in Copper Amine Oxidases by Generating Heterodimers. ChemBioChem 16, 559–564.

Gerstein, M., 2000. Integrative database analysis in structural genomics. Nat Struct Biol 7 Suppl, 960–963.

Gludovacz, E., Maresch, D., Bonta, M., Szöllösi i, H., Furthmüller, P.G., Weik, R., Altmann, F., Limbeck, A., Borth, N., Jilma, B., Boehm, T., 2016. Characterization of recombinant human diamine oxidase (rhDAO) produced in Chinese Hamster Ovary (CHO) cells. J. Biotechnol. 227, 120–130.

Gludovacz, E., Maresch, D., De Carvalho, L.L., Puxbaum, V., Baier, L.J., Sützl, L., Guédez, G., Grünwald-Gruber, C., Ulm, B., Pils, S., Ristl, R., Altmann, F., Jilma, B., Salminen, T.A., Borth, N., Boehm, T., 2018. Oligomannosidic glycans at asn-110 are essential for secretion of human diamine oxidase. J. Biol. Chem. 293, 1070–1087.

Grau-Bové, X., Ruiz-Trillo, I., Rodriguez-Pascual, F., 2015. Origin and evolution of lysyl oxidases. Sci. Rep. 5, 1–11.

Greenwood, F.C., Landon, J., 1966. Growth Hormone Secretion in Response to Stress in Man. Nature 210, 540.

Hagen, J.B., 2000. The origins of bioinformatics. Nat. Rev. Genet. 1, 231–236.

Hirokawa, T., Boon-Chieng, S., Mitaku, S., 1998. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics 14, 378–379.

Holt, A., Smith, D.J., Cendron, L., Zanotti, G., Rigo, A., Di Paolo, M.L., 2008. Multiple binding sites for substrates and modulators of semicarbazide-sensitive amine oxidases: kinetic consequences. Mol. Pharmacol. 73, 525–538.

Huang, Q., Liu, Q., Hao, Q., 2005. Crystal structures of Fms1 and its complex with spermine reveal substrate specificity. J. Mol. Biol. 348, 951–959.

Imamura, Y., Kubota, R., Wang, Y., Asakawa, S., Kudoh, J., Mashima, Y., Oguchi, Y., Shimizu, N., 1997. Human retina-specific amine oxidase (RAO): cDNA cloning, tissue expression, and chromosomal mapping. Genomics 40, 277–283.

Imperiali, B., O'Connor, S.E., 1999. Effect of N-linked glycosylatian on glycopeptide and glycoprotein structure. Curr. Opin. Chem. Biol. 3, 643–649.

Inoue, T., Morita, M., Tojo, T., Nagashima, A., Moritomo, A., Imai, K., Miyake, H., 2013a. Synthesis and SAR study of new thiazole derivatives as vascular adhesion protein-1 (VAP-1) inhibitors for the treatment of diabetic macular edema: part 2. Bioorg. Med. Chem. 21, 2478–94.

Inoue, T., Morita, M., Tojo, T., Nagashima, A., Moritomo, A., Miyake, H., 2013b. Novel 1H-imidazol-2-amine derivatives as potent and orally active vascular adhesion protein-1 (VAP-1) inhibitors for diabetic macular edema treatment. Bioorg. Med. Chem. 21, 3873–81.

Jaakkola, K., Kaunisma, K., Havia, T., Jalkanen, S., Salmi, M., 1999. Human Vascular Adhesion Protein-1 in Smooth Muscle Cells. Am. J. Pathol. 155, 1953–1965.

Jaakkola, K., Nikula, T., Holopainen, R., Vahasilta, T., Matikainen, M.T., Laukkanen, M.L., Huupponen, R., Halkola, L., Nieminen, L., Hiltunen, J., Parviainen, S., Clark, M.R., Knuuti, J., Savunen, T., Kaapa, P., Voipio-Pulkki, L.M., Jalkanen, S., 2000. In vivo detection of vascular adhesion protein-1 in experimental inflammation. Am J Pathol 157, 463–471.

Jakobsson, E., Nilsson, J., Ogg, D., Kleywegt, G.J., 2005. Structure of human

semicarbazide-sensitive amine oxidase/vascular adhesion protein-1. Acta Crystallogr. Sect. D Biol. Crystallogr. 61, 1550–1562.

Jalkanen, S., Karikoski, M., Mercier, N., Koskinen, K., Henttinen, T., Elima, K., Salmivirta, K., Salmi, M., 2007. The oxidase activity of vascular adhesion protein-1 (VAP-1) induces endothelial E- and P-selectins and leukocyte binding. Blood 110, 1864–70.

Jalkanen, S., Salmi, M., 2017. Vascular adhesion protein-1: A cell-surface amine oxidase in translation. Antioxid. Redox Signal. 00, ars.2017.7418.

Janes, S.M., Mu, D., Wemmer, D., Smith, a J., Kaur, S., Maltby, D., Burlingame, a L., Klinman, J.P., 1990. A new redox cofactor in eukaryotic enzymes: 6-hydroxydopa at the active site of bovine serum amine oxidase. Science 248, 981–987.

Jarnicki, A.G., Schilter, H., Liu, G., Wheeldon, K., Essilfie, A.T., Foot, J.S., Yow, T.T., Jarolimek, W., Hansbro, P.M., 2016. The inhibitor of semicarbazide-sensitive amine oxidase, PXS-4728A, ameliorates key features of chronic obstructive pulmonary disease in a mouse model. Br. J. Pharmacol. 173, 3161–3175.

Jensen, S.B., Käkelä, M., Jødal, L., Moisio, O., Alstrup, A.K.O., Jalkanen, S., Roivainen, A., 2017. Exploring the radiosynthesis and in vitro characteristics of [68 Ga]Ga-DOTA-Siglec-9. J. Label. Compd. Radiopharm. 60, 439–449.

Johnson, M.S., Lehtonen, J. V, 2000. Comparision of protein three-dimensional structures, in: Higgins, D., Taylor, W. (Eds.), Bioinformatics: Sequence,Structure and Databanks. Oxford University Press, Oxford, UK, p. 15.

Johnson, M.S., May, A.C.W., Rodinov, M.A., Overington, J.P., 1996. Discrimination of common protein folds: Application of protein structure to sequence/structure comparisons. Methods Enzymol. 266, 575–598.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.

Kaitaniemi, S., Elovaara, H., Grön, K., Kidron, H., Liukkonen, J., Salminen, T., Salmi, M., Jalkanen, S., Elima, K., 2009. The unique substrate specificity of human AOC2, a semicarbazide-sensitive amine oxidase. Cell. Mol. Life Sci. 66, 2743–2757.

Kemik, O., Sümer, A., Kemik, A.S., Itik, V., Dulger, A.C., Purisa, S., Tuzun, S., 2010. Human vascular adhesion proteın-1 (VAP-1): serum levels for hepatocellular carcinoma in non-alcoholic and alcoholic fatty liver disease. World J. Surg. Oncol. 8, 83.

Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B.A., Wang, J., Yu, B., Zhang, J., Bryant, S.H., 2016. PubChem substance and compound databases. Nucleic Acids Res. 44, D1202–D1213.

Kivi, E., Elima, K., Aalto, K., Nymalm, Y., Auvinen, K., Koivunen, E., Otto, D.M., Crocker, P.R., Salminen, T. a, Salmi, M., Jalkanen, S., 2009. Human Siglec-10 can bind to vascular adhesion protein-1 and serves as its substrate. Blood 114, 5385–92.

Klema, V.J., Wilmot, C.M., 2012. The role of protein crystallography in defining the mechanisms of biogenesis and catalysis in copper amine oxidase. Int. J. Mol. Sci. 13, 5375–5405.

Klinman, J.P., 1996. New quinocofactors in eukaryotes. J. Biol. Chem. 271, 27189–27192.

Kumar, S., Stecher, G., Peterson, D., Tamura, K., 2012. MEGA-CC: Computing core of molecular evolutionary genetics analysis program for automated and iterative data analysis. Bioinformatics 28, 2685–2686.

Kurkijärvi, R., Adams, D.H., Leino, R., Möttönen, T., Jalkanen, S., Salmi, M., 1998. Circulating form of human vascular adhesion protein-1 (VAP-1): increased serum levels in inflammatory liver diseases. J. Immunol. 161, 1549–1557.

Lai, J., Jin, J., Kubelka, J., Liberles, D.A., 2012. A phylogenetic analysis of normal modes evolution in enzymes and its relationship to enzyme function. J. Mol. Biol. 422, 442–459.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.

Laymon, C.W., Cumming, H., 1939. Histaminase in the Treatment of Urticaria and Atopic Dermatitis. J. Invest. Dermatol. 2, 301–312.

Le, S.Q., Gascuel, O., 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. 25, 1307–1320.

Lehtonen, J. V, Still, D.-J., Rantanen, V.-V., Ekholm, J., Björklund, D., Iftikhar, Z., Huhtala, M., Repo, S., Jussila, A., Jaakkola, J., Pentikäinen, O., Nyrönen, T., Salminen, T., Gyllenberg, M., Johnson, M.S., 2004. BODIL: a molecular modeling environment for structure-function analysis and drug design. J. Comput. Aided. Mol. Des. 18, 401–419.

Li, H.Y., Jiang, Y. Der, Chang, T.J., Wei, J.N., Lin, M.S., Lin, C.H., Chiang, F.T., Shih, S.R., Hung, C.S., Hua, C.H., Smith, D.J., Vanio, J., Chuang, L.M., 2011. Serum vascular adhesion protein-1 predicts 10-year cardiovascular and cancer mortality in individuals with type 2 diabetes. Diabetes 60, 993–999.

Li, X.-G., Autio, A., Ahtinen, H., Helariutta, K., Liljenbäck, H., Jalkanen, S., Roivainen, A., Airaksinen, A.J., 2013. Translating the concept of peptide labeling with 5-deoxy-5-[18F] fluororibose into preclinical practice: 18F-labeling of Siglec-9 peptide for PET imaging of inflammation. Chem. Commun. 49, 3682–4.

Li, Y.-I., Hung, J.-S., Yu, T.-Y., Liou, J.-M., Wei, J.-N., Kao, H.-L., Chuang, L.-M., Shun, C.-T., Lee, P.-H., Lai, H.-S., Su, C.-Y., Li, H.-Y., Liang, J.-T., 2014. Serum vascular adhesion protein-1 predicts all-cause mortality and cancer-related mortality in subjects with colorectal cancer. Clin. Chim. Acta 428, 51–56.

Liu, T., Kim, D.W., Niitsu, M., Berberich, T., Kusano, T., 2014a. Oryza sativa polyamine oxidase 1 back-converts tetraamines, spermine and thermospermine, to spermidine. Plant Cell Rep. 33, 143–151.

Liu, T., Kim, D.W., Niitsu, M., Maeda, S., Watanabe, M., Kamio, Y., Berberich, T., Kusano, T., 2014b. Polyamine oxidase 7 is a terminal catabolism-type enzyme in oryza sativa and is specifically expressed in anthers. Plant Cell Physiol. 55, 1110–1122.

Lunelli, M., Di Paolo, M.L., Biadene, M., Calderone, V., Battistutta, R., Scarpa, M., Rigo, A., Zanotti, G., 2005. Crystal structure of amine oxidase from bovine serum. J. Mol. Biol. 346, 991–1004.

Luscombe, N.M., Greenbaum, D., Gerstein, M., 2001a. What is bioinformatics? A proposed definition and overview of the field. Methods Inf. Med. 40, 346–358.

Luscombe, N.M., Greenbaum, D., Gerstein, M., 2001b. What is bioinformatics? A proposed definition and overview of the field. Gene Expr. 40, 346–358.

Lyles, G.A., 1996. Mammalian Plasma and Tissue-bond Semicarbazide-sensitive Amine Oxidases: Biochemical, Pharmacological and Toxicological Aspects. Int. J. Biochem. Cell Bio 28, 259–274.

Maintz, L., Novak, N., 2007. Histamine and histamine intolerance. Am. J. Clin. Nutr. 85, 1185–1196.

Maintz, L., Schwarzer, V., Bieber, T., van der

Ven, K., Novak, N., 2008. Effects of histamine and diamine oxidase activities on pregnancy: A critical review. Hum. Reprod.

Mallery, D.L., McEwan, W.A., Bidgood, S.R., Towers, G.J., Johnson, C.M., James, L.C., 2010. Antibodies mediate intracellular immunity through tripartite motif-containing 21 (TRIM21). Proc. Natl. Acad. Sci. 107, 19985 LP-19990.

Maula, S.M., Salminen, T., Kaitaniemi, S., Nymalm, Y., Smith, D.J., Jalkanen, S., 2005. Carbohydrates located on the top of the "cap" contribute to the adhesive and enzymatic functions of vascular adhesion protein-1. Eur. J. Immunol. 35, 2718–2727.

McGrath, A.P., Hilmer, K.M., Collyer, C. a., Shepard, E.M., Elmore, B.O., Brown, D.E., Dooley, D.M., Guss, J.M., 2009. Structure and inhibition of human diamine oxidase. Biochemistry 48, 9810–9822.

McGrath, A.P., Hilmer, K.M., Collyer, C.A., Dooley, D.M., Guss, J.M., 2010. A new crystal form of human diamine oxidase. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 66, 137–142.

Mondovì, B., Agrò, A.F., 1982. Structure and Function of Amine Oxidases, in: Structure and Function Relationships in Biochemical Systems. Advances in Experimental Medicine and Bioligy. pp. 141–153.

Mure, M., 2004. Tyrosine-Derived Quinone Cofactors. Acc. Chem. Res. 37, 131–139.

Mure, M., Mills, S.A., Klinman, J.P., 2002. Catalytic mechanism of the topa quinone containing copper amine oxidases. Biochemistry 41, 9269–9278.

Noonan, T., Lukas, S., Peet, G.W., Pelletier, J., Panzenbeck, M., Hanidu, A., Mazurek, S., Wasti, R., Rybina, I., Roma, T., Kronkaitis, A., Shoultz, A., Souza, D., Jiang, H., Nabozny, G., Modis, L.K., 2013. The oxidase activity of vascular adhesion protein-1 (VAP-1) is essential for function. Am. J. Clin. Exp. Immunol. 2, 172–85.

O'Rourke, A.M., Wang, E.Y., Miller, A., Podar, E.M., Scheyhing, K., Huang, L., Kessler, C., Gao, H., Ton-Nu, H.-T., Macdonald, M.T., Jones, D.S., Linnik, M.D., 2008. Anti-inflammatory effects of LJP 1586 [Z-3-fluoro-2-(4methoxybenzyl) allylamine hydrochloride], an amine-based inhibitor of semicarbazide-sensitive amine oxidase activity. J. Pharmacol. Exp. Ther. 324, 867–75.

Ono, Y., Kim, D.W., Watanabe, K., Sasaki, A., Niitsu, M., Berberich, T., Kusano, T., Takahashi, Y., 2012. Constitutively and highly expressed Oryza sativa polyamine oxidases localize in peroxisomes and catalyze polyamine back conversion. Amino Acids 42, 867–876.

Pannecoeck, R., Serruysa, D., Benmeridjaa, L., Delangheb, J.R., Geelc, N., Speeckaertc, R., Speeckaerta, M.M., 2015. Vascular adhesion protein-1: Role in human pathology and application as a biomarker. Crit. Rev. Clin. Laboraty Sci. 8363, 1–17.

Parsons, M.R., Convery, M.A., Wilmot, C.M., Yadav, K.D.S., Blakeley, V., Corner, A.S., Phillips, S., McPherson, M.J., Knowles, P.F., 1995. Crystal structure of a quinoenzyme: copper amine oxidase of Escherichia coli at 2 Å resolution. Structure 3, 1171–1184.

Petersen, T.N., Brunak, S., Heijne, G. von, Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods 8, 785–786.

Rajtar, S., Irman-Florjanc, T., 2007. Amitriptyline affects histamine-N-methyltransferase and diamine oxidase activity in rats and guinea pigs. Eur. J. Pharmacol. 574, 201–208.

Repessé, X., Moldes, M., Muscat, A., Vatier, C., Chetrite, G., Gille, T., Planes, C., Filip, A., Mercier, N., Duranteau, J., Fève, B., 2015. Hypoxia inhibits semicarbazide-

sensitive amine oxidase activity in adipocytes. Mol. Cell. Endocrinol. 411, 58–66.

Saitou, N., Nei, M., 1987. The neighbour-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evo 4, 406–425.

Salentin, S., Schreiber, S., Haupt, V.J., Adasme, M.F., Schroeder, M., 2015. PLIP: Fully automated protein-ligand interaction profiler. Nucleic Acids Res. 43, W443–W447.

Sali, A., Blundell, T., 1993. Comparative Protein Modelling by Satisfaction of Spatial Restraints. J. Mol. Biol. 234, 779–815.

Salmi, M., Jalkanen, S., 2014. Ectoenzymes in leukocyte migration and their therapeutic potential. Semin. Immunopathol. 36, 163–176.

Salmi, M., Jalkanen, S., 2011. Homing-associated molecules CD73 and VAP-1 as targets to prevent harmful inflammations and cancer spread. FEBS Lett. 585, 1543–50.

Salmi, M., Jalkanen, S., 2001. VAP-1: an adhesin and an enzyme. Trends Immunol. - 22, 211–216.

Salmi, M., Jalkanen, S., 1996. Human vascular adhesion protein 1 (VAP-1) is a unique sialoglycoprotein that mediates carbohydrate-dependent binding of lymphocytes to endothelial cells. J. Exp. Med.183,569–79.

Salmi, Marko, Kalimo, K., Jalkanen, S., 1993. Induction and Function of Vascular Adhesion Protein-1 at Sites of Inflammation. J. Exp. Med. 178, 2255–2260.

Salminen, T. a., Smith, D.J., Jalkanen, S., Johnson, M.S., 1998. Structural model of the catalytic domain of an enzyme with cell adhesion activity: human vascular adhesion protein-1 (HVAP-1) D4 domain is an amine oxidase. Protein Eng. Des. Sel. 11, 1195–1204.

Sattler, J., Hesterberg, R., Lorenz, W., Schmidt, U., Crombach, M., Stahlknecht, C.D., 1985. Inhibition of human and canine diamine oxidase by drugs used in an intensive care unit: relevance for clinical side effects? Agents Actions 16, 91–94.

Schilter, H.C., Collison, A., Russo, R.C., Foot, J.S., Yow, T.T., Vieira, A.T., Tavares, L.D., Mattes, J., Teixeira, M.M., Jarolimek, W., 2015. Effects of an anti-inflammatory VAP-1/SSAO inhibitor, PXS-4728A, on pulmonary neutrophil migration. Respir. Res. 16, 1–14.

Schwelberger, H.G., 2010. Structural organization of mammalian copper-containing amine oxidase genes. Inflamm. Res. 59 Suppl 2, S223-5.

Schwelberger, H.G., 2007. The origin of mammalian plasma amine oxidases. J. Neural Transm. 114, 757–62.

Schwelberger, H.G., 2006. Origins of plasma amine oxidases in different mammalian species. Inflamm. Res. 55 Suppl 1, S57-8.

Schwelberger, H.G., Bodner, E., 1999. Analysis of the glycosylation of porcine diamine oxidase. Inflamm. Res. 48 Suppl 1, S83–S84.

Shen, S.H., Wertz, D.L., Klinman, J.P., 2012. Implication for functions of the ectopic adipocyte copper amine oxidase (AOC3) from purified enzyme and cell-based kinetic studies. PLoS One 7, 1–11.

Soltis, D.E., Soltis, P.S., 2003. Applying the Bootstrap in Phylogeny Reconstruction. Stat. Sci. 18, 256–267.

Sonnhammer, E.L., von Heijne, G., Krogh, a, 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 175–182.

Stolen, C.M., Yegutkin, G.G., Kurkijärvi, R., Bono, P., Alitalo, K., Jalkanen, S., 2004. Origins of serum semicarbazide-sensitive amine oxidase. Circ. Res. 95, 50–7.

Takahashi, Y., Cong, R., Sagor, G.H.M., Niitsu, M., Berberich, T., Kusano, T., 2010. Characterization of five polyamine oxidase isoforms in Arabidopsis thaliana. Plant Cell Rep. 29, 955–965.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tannous, A., Pisoni, G.B., Hebert, D.N., Molinari, M., 2015. N-linked sugar-regulated protein folding and quality control in the ER. Semin. Cell Dev. Biol. 41,79–89.

Tavladoraki, P., Cervelli, M., Antonangeli, F., Minervini, G., Stano, P., Federico, R., Mariottini, P., Polticelli, F., 2011. Probing mammalian spermine oxidase enzyme-substrate complex through molecular modeling, site-directed mutagenesis and biochemical characterization. Amino Acids 40, 1115–1126.

Tavladoraki, P., Cona, A., Angelini, R., 2016. Copper-containing amine oxidases and FAD-dependent polyamine oxidases are key players in plant tissue differentiation and organ development. Front. Plant Sci. 7, 824.

Tavladoraki, P., Rossi, M.N., Saccuti, G., Perez-Amador, M.A., Polticelli, F., Angelini, R., Federico, R., 2006. Heterologous expression and biochemical characterization of a polyamine oxidase from Arabidopsis involved in polyamine back conversion. Plant Physiol. 141, 1519–32.

Tavladoraki, P., Schininà, M.E., Cecconi, F., Di Agostino, S., Manera, F., Rea, G., Mariottini, P., Federico, R., Angelini, R., 1998. Maize polyamine oxidase: Primary structure from protein and cDNA sequencing. FEBS Lett. 426, 62–66.

Toninello, A., Pietrangeli, P., De Marchi, U., Salvi, M., Mondovì, B., 2006. Amine oxidases in apoptosis and cancer. Biochim. Biophys. Acta - Rev. Cancer 1765,1–13.

UniProt: the universal protein knowledgebase, 2017. . Nucleic Acids Res. 45, D158–D169.

Vainio, P.J., Kortekangas-savolainen, O., Mikkola, J.H., Jaakkola, K., Kalimo, K., Jalkanen, S., Veromaa, T., 2005. Safety of Blocking Vascular Adhesion Protein-1 in Patients with Contact Dermatitis 429–435.

Venkatesh, B., Lee, A.P., Ravi, V., Maurya, A.K., Lian, M.M., Swann, J.B., Ohta, Y., Flajnik, M.F., Sutoh, Y., Kasahara, M., Hoon, S., Gangu, V., Roy, S.W., Irimia, M., Korzh, V., Kondrychyn, I., Lim, Z.W., Tay, B.H., Tohari, S., Kong, K.W., Ho, S., Lorente-Galdos, B., Quilez, J., Marques-Bonet, T., Raney, B.J., Ingham, P.W., Tay, A., Hillier, L.W., Minx, P., Boehm, T., Wilson, R.K., Brenner, S., Warren, W.C., 2014. Elephant shark genome provides unique insights into gnathostome evolution. Nature 505, 174–179.

Virtanen, H., Autio, A., Siitonen, R., Liljenbäck, H., Saanijoki, T., Lankinen, P., Mäkilä, J., Käkelä, M., Teuho, J., Savisto, N., Jaakkola, K., Jalkanen, S., Roivainen, A., 2015. 68Ga-DOTA-Siglec-9--a new imaging tool to detect synovitis. Arthritis Res. Ther. 17, 308.

Wallner, B., Elofsson, A., 2003. Can correct protein models be identified? Protein Sci. 12, 1073–86.

Wheeler, T.J., Clements, J., Finn, R.D., 2014. Skylign: a tool for creating informative, interactive logos representing sequence alignments and profile hidden Markov models. BMC Bioinformatics 15, 7.

Wiederstein, M., Sippl, M.J., 2007. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 35.

Wilce, M.C.J., Dooley, D.M., Freeman, H.C., Guss, J.M., Matsunami, H., McIntire, W.S., Ruggiero, C.E., Tanizawa, K., Yamaguchi, H., 1997. Crystal structures of the copper-containing amine oxidase

from Arthrobacter globiformis in the holo and apo forms: Implications for the biogenesis of topaquinone. Biochemistry 36, 16116–16133.

Yamada, M., Yasuhara, H., 2004. Clinical Pharmacology of MAO Inhibitors: Safety and Future. Neurotoxicology 25, 215–221.

Yamaki, S., Koga, Y., Nagashima, A., Kondo, M., Shimada, Y., Kadono, K., Moritomo, A., Yoshihara, K., 2017a. Synthesis and pharmacological evaluation of glycine amide derivatives as novel vascular adhesion protein-1 inhibitors without CYP3A4 and CYP2C19 inhibition. Bioorganic Med. Chem. 25, 4110–4122.

Yamaki, S., Suzuki, D., Fujiyasu, J., Neya, M., Nagashima, A., Kondo, M., Akabane, T., Kadono, K., Moritomo, A., Yoshihara, K., 2017b. Bioorganic & Medicinal Chemistry Synthesis and structure activity relationships of glycine amide derivatives as novel Vascular Adhesion Protein-1 inhibitors. Bioorg. Med. Chem. 25, 187–201.

Yamaki, S., Yamada, H., Nagashima, A., Kondo, M., Shimada, Y., Kadono, K., Yoshihara, K., 2017c. Synthesis and structure activity relationships of carbamimidoylcarbamate derivatives as novel vascular adhesion protein-1 inhibitors. Bioorganic Med. Chem. 25, 6024–6038.

Yasuda, H., Toiyama, Y., Ohi, M., Mohri, Y., Miki, C., Kusunoki, M., 2011. Serum soluble vascular adhesion protein-1 is a valuable prognostic marker in gastric cancer. J. Surg. Oncol. 103, 695–699.

Zhang, Q., Mashima, Y., Noda, S., Imamura, Y., Kudoh, J., Shimizu, N., Nishiyama, T., Umeda, S., Oguchi, Y., Tanaka, Y., Iwata, T., 2003. Characterization of AOC2 gene encoding a copper-binding amine oxidase expressed specifically in retina. Gene 318, 45–53.

Zorzano, A., Abella, A., Marti, L., Carpéné, C., Palacín, M., Testar, X., 2003. Semicarbazide-sensitive amine oxidase activity exerts insulin-like effects on glucose metabolism and insulin-signaling pathways in adipose cells. Biochim. Biophys. Acta - Proteins Proteomics 1647, 3–9.

Zuckerkandl, E., Pauling, L., 1965. Molecules as documents of evolutionary history. J. Theor. Biol. 8, 357–366.

ISBN 978-952-12-3763-8